

(FILE 'USPAT' ENTERED AT 10:29:32 ON 30 MAY 96)

L1 0 S MEGAKARYOCYTE(W) KINASE  
L2 0 S MEGAKARYOCYTE(5A) KINASE  
L3 3 S MEGAKARYOCYTE(P) KINASE

=> d 1-3

1. 5,491,242, Feb. 13, 1996, Protein kinase C inhibitors; James R. Gillig, et al., 548/455 [IMAGE AVAILABLE]

2. 5,481,003, Jan. 2, 1996, Protein kinase C inhibitors; James R. Gillig, et al., 548/455, 312.1 [IMAGE AVAILABLE]

3. 5,378,464, Jan. 3, 1995, Modulation of inflammatory responses by administration of GMP-140 or antibody to GMP-140; Rodger P. McEver, 424/143.1; 514/8 [IMAGE AVAILABLE]

L4 0 Non-receptor (5A) Tyrosine<sup>(w)</sup> Kinase

Set	Items	Description
S1	0	MEGAKARYOCYTE (W) KINASE
S2	5	MEGAKARYOCYTE (3N) KINASE
S3	67	(NON (W) RECEPTOR) (5N) (TYROSINE (W) KINASE)
S4	1	S3 AND MEGAKARYOCYTE

?d s2/3/all

Display 2/3/1

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09606415 96128015

Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*.

Ems SC; Morden CW; Dixon CK; Wolfe KH; dePamphilis CW; Palmer JD

Department of Biology, Indiana University, Bloomington 47405, USA.

Plant Mol Biol (NETHERLANDS) Nov 1995, 29 (4) p721-33, ISSN 0167-4412

Journal Code: A60

Contract/Grant No.: GM-35087, GM, NIGMS

Languages: ENGLISH

Document type: JOURNAL ARTICLE

- end of record -

?

Display 2/3/2

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09308492 95238492

The MATK tyrosine kinase interacts in a specific and SH2-dependent manner with c-Kit.

Jhun BH; Rivnay B; Price D; Avraham H

Department of Medicine, Deaconess Hospital, Harvard Medical School, Boston, Massachusetts 02215, USA.

J Biol Chem (UNITED STATES) Apr 21 1995, 270 (16) p9661-6, ISSN 0021-9258 Journal Code: HIV

Contract/Grant No.: HL51456, HL, NHLBI; HL46668, HL, NHLBI

Languages: ENGLISH

Document type: JOURNAL ARTICLE

- end of record -

?

Display 2/3/3

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09200565 95130565

Structural and functional studies of the intracellular tyrosine kinase  
MATK gene and its translated product.

Avraham S; Jiang S; Ota S; Fu Y; Deng B; Dowler LL; White RA; Avraham H  
Division of Hematology/Oncology, New England Deaconess Hospital, Harvard  
Medical School, Boston, Massachusetts 02215.

J Biol Chem (UNITED STATES) Jan 27 1995, 270 (4) p1833-42, ISSN  
0021-9258 Journal Code: HIV

Contract/Grant No.: R01 HL51456, HL, NHLBI; R01 HL46668, HL, NHLBI

Languages: ENGLISH

Document type: JOURNAL ARTICLE

- end of record -

?

Display 2/3/4

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09176341 95106341

Identification and characterization of Batk, a predominantly  
brain-specific non-receptor protein tyrosine kinase related to Csk.

Kuo SS; Moran P; Gripp J; Armanini M; Phillips HS; Goddard A; Caras IW  
Department of Neurobiology, Genentech, Inc., South San Francisco,  
California 94080.

J Neurosci Res (UNITED STATES) Aug 15 1994, 38 (6) p705-15, ISSN  
0360-4012 Journal Code: KAC

Languages: ENGLISH

Document type: JOURNAL ARTICLE

- end of record -

?

Display 2/3/5

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1996 Knight-Ridder Info. All rts. reserv.

08802408 94117408

Identification and characterization of a novel tyrosine kinase from  
megakaryocytes.

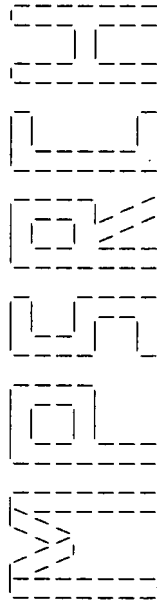
Bennett BD; Cowley S; Jiang S; London R; Deng B; Grabarek J; Groopman JE;  
Goeddel DV; Avraham H

Division of Hematology/Oncology, New England Deaconess Hospital, Harvard  
Medical School, Boston, Massachusetts 02215.

J Biol Chem (UNITED STATES) Jan 14 1994, 269 (2) p1068-74, ISSN  
0021-9258 Journal Code: HIV  
Contract/Grant No.: HL33774, HL, NHLBI; HL42112, HL, NHLBI; HL43510, HL,  
NHLBI; +  
Languages: ENGLISH  
Document type: JOURNAL ARTICLE

- end of display -

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri May 31 09:32:36 1996; MasPar time 21.94 Seconds  
588,464 Million cell updates/sec

```
>US-08-426-509-6
Description: (1-511) from US08426509.pep
Perfect Score: 3821
Sequence: 1 MVSVCMSTICQRLWEYLEPY.....
```

Scoring table: PAM 150  
Gap 11

Searched: 82306 seqs. 25270970 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

```
Database:
pir46
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3 7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2
```

Statistics: Mean 48.603; Variance 117.818; scale 0.413

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2032	53.2	505	4	S24550	protein-tyrosine kin	0.00e+00	
2	2026	53.0	506	4	S24553	protein-tyrosine kin	0.00e+00	
3	1801	47.1	509	1	TWHA5T	protein-tyrosine kin	6.43e-281	
4	1754	45.9	536	10	S33569	protein-tyrosine kin	1.02e-272	
5	1754	45.9	536	4	S29626	protein kinase yrk (	1.02e-272	
6	1745	45.7	362	4	S24551	protein-tyrosine kin	3.80e-271	
7	1734	45.4	537	1	TWHSY	protein-tyrosine kin	3.15e-269	
8	1730	45.3	537	4	A43806	protein-tyrosine kin	1.57e-268	
9	1726	45.2	537	1	TWHSR	protein-tyrosine kin	7.82e-268	
10	1685	44.1	512	4	A39719	protein-tyrosine kin	1.10e-260	
11	1680	44.0	512	1	TWHD1Y	protein-tyrosine kin	8.15e-260	
12	1682	44.0	532	4	B34104	protein-tyrosine kin	3.65e-260	

13	1673	43.8	532	4	A34104	protein-tyrosine	kin	1.35e-255
14	1674	43.8	533	1	TVCHS	protein-tyrosine	kin	9.05e-259
15	1664	43.5	507	4	A39399	protein-tyrosine	kin	5.00e-257
16	1662	43.5	587	1	TVFVPR	protein-tyrosine	kin	1.11e-256
17	1660	43.4	517	4	TVH0UC	protein-tyrosine	kin	2.49e-256
18	1659	43.4	505	4	A43807	protein-tyrosine	kin	3.71e-256
19	1659	43.4	517	11	S24947	protein-tyrosine	kin	3.71e-256
20	1657	43.4	568	1	TVFVS1	protein-tyrosine	kin	8.28e-256
21	1652	43.2	557	1	TVFVS9	protein-tyrosine	kin	6.15e-255
22	1647	43.1	509	4	A23639	protein-tyrosine	kin	4.51e-254
23	1645	43.1	529	1	TVHFUR	protein-tyrosine	kin	1.07e-253
24	1646	43.1	542	1	TVHUSC	protein-tyrosine	kin	6.82e-254
25	1644	43.0	541	4	A43610	protein-tyrosine	kin	1.52e-253
26	1639	42.9	503	1	TVMSHC	protein-tyrosine	kin	1.13e-252
27	1639	42.9	503	4	A39973	protein-tyrosine	kin	1.13e-252
28	1640	42.9	537	4	A45501	protein-tyrosine	kin	7.56e-253
29	1638	42.9	541	11	S31645	protein-tyrosine	kin	1.69e-252
30	1635	42.8	503	4	JO1321	protein-tyrosine	kin	5.61e-252
31	1636	42.8	541	1	TVCHS	protein-tyrosine	kin	3.76e-252
32	1632	42.8	543	1	TVHUYD	protein-tyrosine	kin	5.61e-252
33	1627	42.6	509	1	OKHULK	protein-tyrosine	kin	1.39e-250
34	1625	42.5	503	11	S18974	protein-tyrosine	kin	3.09e-250
35	1622	42.4	526	1	TVFV60	protein-tyrosine	kin	1.03e-249
36	1617	42.3	534	4	A44991	protein-tyrosine	kin	7.63e-249
37	1610	42.1	528	1	TVFV69	protein-tyrosine	kin	1.26e-247
38	1605	42.0	499	4	A40092	protein-tyrosine	kin	9.35e-247
39	1598	41.8	505	11	S51647	protein-tyrosine	kin	1.54e-245
40	1596	41.8	534	4	S33568	protein-tyrosine	kin	3.44e-245
41	1591	41.6	526	1	TVFVPR	protein-tyrosine	kin	2.55e-244
42	1589	41.6	526	4	S15582	protein-tyrosine	kin	5.68e-244
43	1589	41.6	545	7	S52313	pp62v protein - Rous	5.68e-244	
44	1588	41.6	546	7	S52314	pp62v protein - Rous	8.48e-244	
45	1584	41.5	526	13	S32774	protein-tyrosine	kin	4.21e-243

## ALIGNMENTS

```

1 RESULT
ENTRY protein-tyrosine kinase (EC 2.7.1.112) 1 - freshwater sponge
TITLE (Spongilla lacustris)
ALTERNATE_NAMES src-type tyrosine kinase 1
ORGANISM #formal_name Spongilla lacustris
DATE 07-May-1993 #sequence_revision 07-May-1993 #text_change
28-Oct-1994
ACCESSIONS S24550
REFERENCE S24550
AUTHORS Raulf, F.
SUBMISSION submitted to the EMBL Data Library, September 1991
 accession S24550
#molecule_type mRNA
#residues 1-505 #label RAU
#cross-references EMBL:X61601
GENETICS
#gene
CLASSIFICATION #superfamily protein-tyrosine kinase src; protein kinase
ATP; phosphotransferase; tyrosine-specific protein kinase
KEYWORDS #domain SH3 homology #label SH3\
61-111 #domain SH2 homology #label SH2\
122-214 #domain protein kinase homology #label KIN\
238-496 #region nucleotide-binding motif\
246-254 #active site Lys #status predicted
FEATURE
268

```











May 31 09:27

US-08-426-569-6.rag

3

CC activity. 505 AA;  
 SQ Sequence 505 AA;  
 Query Match 97.7%; Score 3734; DB 14; Length 505;  
 Best Local Similarity 99.6%; Pred. No. 0.00e+00;  
 Matches 504; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 Db 1 manicqrlwleypclstlstdkstvienpgalcspsqqrhghyvalfdyqartaed 60  
 QY 6 MSNIGQRLMEYLEPILPCSLSTLSTADKSTVIENPGALCSPSQQRHGHYFVALFDYQARTAD 65  
 Db 61 lafragdk-lqvldt.lhegwfarhlekrrdsgsqglqgypipnyvaedrsdqaeppwffg 119  
 QY 66 LSFRAGDKKLQVLDLHEGWWFARHLEKRRDSSQQLQGYIPSNVVAEDRSIQAEPPWFFG 125  
 Db 120 algrsdaekqlysenktgsfliresesqgfeislvdgavvkhryikrlddegffltr 179  
 QY 126 AIGRSDAEKQLLYSENKTSGLFIRESESQKGEFSLSLVDGAVVVKHYRIKRLDEGGFFLTR 185  
 Db 180 rriifstinefsvhytktsdglcvkigkpcikiqvpapfdlsyktvdqweidrnslqlkr 239  
 QY 186 RRIIFSTINEFVSHYTKTSGLCVKLGKPCIKIQVPAPFDLSYKTVQDWEIDRNSIQLLKR 245  
 Db 240 lsgsqfgevewglwnnttpvavktlkpgsmdpndflreaqimknhrhpkliqlayavctle 299  
 QY 246 LSGSQFGEVWEGWLNNTTPVAVKTLKPGSMDPNDFLREAQIMKNLRHPKLIQLYAVCTLE 305  
 Db 300 dpyiitelmrhgsiqeylqndtgskihltqqvmaaqvasgmaylesrnyihrdlaarn 359  
 QY 306 DPYIITELMRHGSIQEYLQNDTGSKIHLLTQQYDMAAQVASGMAYLESRYNIHRDLAARN 365  
 Db 360 vlvghehniykvadfglarvfkvdnediyesrheiklpvktapeairnsksdwsf 419  
 QY 366 VLVGHEHNIYKVADFGLARVFKVDNEDIYESRHEIKLPVKTAPEAIRNSKFSIKSDWSF 425  
 Db 420 gillyeiitygmypsgmtgaqviqlaqnyrlpqsncpqgfymilecwnaepkerpt 479  
 QY 426 GILLYEIITYGKMPYSGMTGAQVIQMLAQNYRLPQSNCPQGFYIMILECWNAPKERPT 485  
 Db 480 fetlwkledyfetdsysdannfir 505  
 QY 486 FETLWKLEDYFETDSSYSANNFIR 511

## RESULT 2

ID R41941 standard; Protein; 505 AA.  
 AC R41941;  
 DT 10-MAR-1994 (first entry)  
 DE pTK gene LptK-2 prod.  
 KW pTK; protein tyrosine kinase; catalytic domain; c-kit; megakaryocyte;  
 KW lymphocyte; amplification; primer; polymerase chain reaction; PCR.  
 OS Homo sapiens.  
 PN W09315201-A.  
 PD 05-AUG-1993.  
 PF 22-JAN-1993; U00586.  
 PR 22-JAN-1992; US-826935.  
 PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.  
 PI Abraham H, Cowley S, Groopman J, Scadden D;  
 DR WPI; 93-320330/40.  
 DR N-PSDB; Q49754.  
 PT New protein tyrosine kinase genes and proteins encoded by genes -  
 PT are of human mega-karyocytic origin  
 PS Claim 3; Fig 5; 60pp; English.  
 CC pTK genes were identified using two sets of degenerative

May 31 09:27

US-08-426-509-6.rag

4

CC oligonucleotide primers: a first set which amplifies all pTK DNA  
 CC segments (Q49743-44), and a second set which amplifies highly  
 CC conserved sequences present in the catalytic domain of the c-kit  
 CC subgroup of pTKs (Q49745-46). The pTK genes identified are described  
 CC in Q49747-57 and R41897-02.  
 CC The LptKs are expressed in lymphocytic cells, as well as  
 CC megakaryocytic cells. The partial and full-length LptK2 gene  
 CC sequences are given in Q49749 and Q49754 respectively. The  
 CC protein sequence corresp. to Q49749 is claimed (claim 7) and  
 CC stated as given in the specification, however is missing from  
 CC the publication.  
 SQ Sequence 505 AA;  
 Query Match 97.7%; Score 3734; DB 8; Length 505;  
 Best Local Similarity 99.6%; Pred. No. 0.00e+00;  
 Matches 504; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 Db 1 manicqrlwleypclstlstdkstvienpgalcspsqqrhghyvalfdyqartaed 60  
 QY 6 MSNIGQRLMEYLEPILPCSLSTLSTADKSTVIENPGALCSPSQQRHGHYFVALFDYQARTAD 65  
 Db 61 lafragdk-lqvldt.lhegwfarhlekrrdsgsqglqgypipnyvaedrsdqaeppwffg 119  
 QY 66 LSFRAGDKKLQVLDLHEGWWFARHLEKRRDSSQQLQGYIPSNVVAEDRSIQAEPPWFFG 125  
 Db 120 algrsdaekqlysenktgsfliresesqgfeislvdgavvkhryikrlddegffltr 179  
 QY 126 AIGRSDAEKQLLYSENKTSGLFIRESESQKGEFSLSLVDGAVVVKHYRIKRLDEGGFFLTR 185  
 Db 180 rriifstinefsvhytktsdglcvkigkpcikiqvpapfdlsyktvdqweidrnslqlkr 239  
 QY 186 RRIIFSTINEFVSHYTKTSGLCVKLGKPCIKIQVPAPFDLSYKTVQDWEIDRNSIQLLKR 245  
 Db 240 lsgsqfgevewglwnnttpvavktlkpgsmdpndflreaqimknhrhpkliqlayavctle 299  
 QY 246 LSGSQFGEVWEGWLNNTTPVAVKTLKPGSMDPNDFLREAQIMKNLRHPKLIQLYAVCTLE 305  
 Db 300 dpyiitelmrhgsiqeylqndtgskihltqqvmaaqvasgmaylesrnyihrdlaarn 359  
 QY 306 DPYIITELMRHGSIQEYLQNDTGSKIHLLTQQYDMAAQVASGMAYLESRYNIHRDLAARN 365  
 Db 360 vlvghehniykvadfglarvfkvdnediyesrheiklpvktapeairnsksdwsf 419  
 QY 366 VLVGHEHNIYKVADFGLARVFKVDNEDIYESRHEIKLPVKTAPEAIRNSKFSIKSDWSF 425  
 Db 420 gillyeiitygmypsgmtgaqviqlaqnyrlpqsncpqgfymilecwnaepkerpt 479  
 QY 426 GILLYEIITYGKMPYSGMTGAQVIQMLAQNYRLPQSNCPQGFYIMILECWNAPKERPT 485  
 Db 480 fetlwkledyfetdsysdannfir 505  
 QY 486 FETLWKLEDYFETDSSYSANNFIR 511  
 RESULT 3  
 ID R39706 standard; Protein; 536 AA.  
 AC R39706;  
 DT 23-DEC-1993 (first entry)  
 DE Human pp60 c-src protein.  
 KW Endothelial; tyrosine kinase protein; pp60 c-src; ss.  
 OS Homo sapiens.  
 PN W09314193-A.  
 PD 22-JUL-1993.  
 PF 05-JAN-1993; US00445.

May 31 09:27

US-08-426-509-6.rag

5

PR 06-JAN-1992; US-820011.  
 PA (UYA ) UNIV YALE.  
 PI Bell L, Luthringer DJ, Madri JA, Warren SL;  
 DR WP1; 93-243209/30.  
 DR P-PSDB; R39705.  
 PT Genetically engineered endothelial cells - which exhibit enhanced  
 PT cell migration, urokinase-type plasminogen activator activity,  
 PT and reduced mononuclear cell adhesion and fibronectin prodn  
 PS Disclosure; Page 75-77; 91pp; English.  
 CC The DNA encoding a portion or (more preferably) the entire pp60  
 CC c-src polypeptide (Given in Q4688) is used to transform endothelial  
 CC cells. Transformed cells produce increased amounts of pp60 c-src and  
 CC have improved therapeutic properties. They migrate at faster rates  
 CC than non-transformed counterparts; have an enhanced ability to  
 CC inhibit the formation of thrombi and/or dissolve thrombi once they  
 CC have formed and exhibit reduced mononuclear cell adhesion. They can  
 CC also be used to improve the success of surgical procedures such as  
 CC coronary angioplasty, heart bypass surgery, vessel graft and stent  
 CC implantation.  
 SQ Sequence 536 AA;

Query Match 43.9%; Score 1679; DB 8; Length 536;  
 Best Local Similarity 51.3%; Pred. No. 2,18e-143;  
 Matches 238; Conservative 97; Mismatches 107; Indels 22; Gaps 14;

Db 89 fvalydesrtetdlsfkqer-lqivntegdwlahtst---q--q-tgyipsnyv 140  
 QY 52 FVALFDYQARTAEIDLSFRAGDKKQLVDTLHEGWMFAHLEKRRDGSQQIQGYIPSNYV 111  
 Db 141 apsdaiqaeevfgkitrreserlllnaenprgtflvresettkayclsvsfdnagl 200  
 QY 112 AEDRSIAQAEFPWFCAIGRSDAEKQLLYSENKTCGSLFIRESEKQGEFSLV--LDGA--V 167  
 Db 201 nvkhykirkldsggyfytstqfnslqqlvayvskhadgichrlttvcpstsk-pqtqqla 259  
 QY 168 -VKHYRKRILDEGGFFLTRRRIFSTNEFVSHYTKTSGLCVKLGKPCIKIQVPAPFDLS 226  
 Db 260 -k--daweipresrlrlevkkgqcgfgevmgtwtgtrvaiktikpgtmspeafleaqv 316  
 QY 227 YKTVDQWEIDRNSIQLKRLGSGQFGEVWGLNNNTPVAVKTKPGSMDPNDFLREAQI 286  
 Db 317 mkkrlrheklvqlayvvs-eepiyivtymksqslldfkgemgkyrlrplqldvmaaqvas 375  
 QY 287 MKNLRHFKLIQLYAVCTLEDPIYIITELMRHGSIQEYLQNDTGSKIHLTQQYDMAAQVAS 346  
 Db 376 gmayvermyvhrdilaanilvgenlvckvadfglarl--iedne-ytargakfpikwt 432  
 QY 347 GMAYLESRNYIHRDLAARNVIVGEHNIYKVADFGLARVFKVDNEDIYESRHEIKLPVKWT 406  
 Db 433 apeaalygrftiksdwsgflltelttkgrvypgmvrnreldqvergymppcepe 492  
 QY 407 APEAIRSNKFSIKSDWWSFGILLIYITGKMPYSGMTGAQVQIOMLAQNYRLPQPSNCPQ 466  
 Db 493 slhldmcqwrkdeperpfeylqafledyftstepyqpgenl 536  
 QY 467 QFYNTMLECNAPKERPFTETLRWKLEDYF-ETDSSYSYDANF 509

RESULT 4  
 ID R39705 standard; Protein; 533 AA.  
 AC R39705;  
 DT 23-DEC-1993 (first entry)  
 DE Chicken pp60 c-src protein.  
 KW Endothelial; tyrosine kinase protein; pp60 c-src; ss.

May 31 09:27

US-08-426-509-6.rag

6

OS Gallus gallus.  
 PN W09314193-A.  
 PD 22-JUL-1993.  
 PF 05-JAN-1993; US00445.  
 PR 06-JAN-1992; US-820011.  
 PA (UYA ) UNIV YALE.  
 PI Bell L, Luthringer DJ, Madri JA, Warren SL;  
 DR WP1; 93-243209/30.  
 DR P-PSDB; R39705.  
 PT Genetically engineered endothelial cells - which exhibit enhanced  
 PT cell migration, urokinase-type plasminogen activator activity,  
 PT and reduced mononuclear cell adhesion and fibronectin prodn  
 PS Disclosure; Page 64-66; 91pp; English.  
 CC The DNA encoding a portion or (more preferably) the entire pp60  
 CC c-src polypeptide (Given in Q4687) is used to transform endothelial  
 CC cells. Transformed cells produce increased amounts of pp60 c-src and  
 CC have improved therapeutic properties. They migrate at faster rates  
 CC than non-transformed counterparts; have an enhanced ability to  
 CC inhibit the formation of thrombi and/or dissolve thrombi once they  
 CC have formed and exhibit reduced mononuclear cell adhesion. They can  
 CC also be used to improve the success of surgical procedures such as  
 CC coronary angioplasty, heart bypass surgery, vessel graft and stent  
 CC implantation.  
 SQ Sequence 533 AA;

Query Match 43.8%; Score 1674; DB 8; Length 533;  
 Best Local Similarity 51.1%; Pred. No. 6.48e-143;  
 Matches 237; Conservative 98; Mismatches 107; Indels 22; Gaps 15;

Db 86 fvalydesrtetdlsfkqer-lqivntegdwlahtst-tg---q-tgyipsnyv 137  
 QY 52 FVALFDYQARTAEIDLSFRAGDKKQLVDTLHEGWMFAHLEKRRDGSQQIQGYIPSNYV 111  
 Db 138 apsdaiqaeevfgkitrreserlllnaenprgtflvresettkayclsvsfdnagl 197  
 QY 112 AEDRSIAQAEFPWFCAIGRSDAEKQLLYSENKTCGSLFIRESEKQGEFSLV--LDGA--V 167  
 Db 198 nvkhykirkldsggyfytstqfnslqqlvayvskhadgichrlttvcpstsk-pqtqqla 256  
 QY 168 -VKHYRKRILDEGGFFLTRRRIFSTNEFVSHYTKTSGLCVKLGKPCIKIQVPAPFDLS 226  
 Db 257 -k--daweipresrlrlevkkgqcgfgevmgtwtgtrvaiktikpgtmspeafleaqv 313  
 QY 227 YKTVDQWEIDRNSIQLKRLGSGQFGEVWGLNNNTPVAVKTKPGSMDPNDFLREAQI 286  
 Db 314 mkkrlrheklvqlayvvs-eepiyivtymksqslldfkgemgkyrlrplqldvmaaqvas 372  
 QY 287 MKNLRHFKLIQLYAVCTLEDPIYIITELMRHGSIQEYLQNDTGSKIHLTQQYDMAAQVAS 346  
 Db 373 gmayvermyvhrdilaanilvgenlvckvadfglarl--iedne-ytargakfpikwt 429  
 QY 347 GMAYLESRNYIHRDLAARNVIVGEHNIYKVADFGLARVFKVDNEDIYESRHEIKLPVKWT 406  
 Db 430 apeaalygrftiksdwsgflltelttkgrvypgmvrnreldqvergymppcepe 489  
 QY 407 APEAIRSNKFSIKSDWWSFGILLIYITGKMPYSGMTGAQVQIOMLAQNYRLPQPSNCPQ 466  
 Db 490 slhldmcqwrkdeperpfeylqafledyftstepyqpgenl 533  
 QY 467 QFYNTMLECNAPKERPFTETLRWKLEDYF-ETDSSYSYDANF 509

RESULT 5  
 ID R14201 standard; Protein; 417 AA.

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

(TM)

\*\*\*\*\*

MPsrch\_mn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Fri May 31 11:36:53 1996; MasPar time 1686.87 Seconds  
Tabular output not generated. 1162.552 Million cell updates/sec

Title: &gt;US-08-426-509-5

Description: (1-2770) from US08426509.seq

Perfect Score: 2770

N.A. Sequence: 1 CCGGACTGGTCCGAGACAG.....AGTACCAAAAAAAAAA 2770

Comp: GGCGCTGACGAGCTTCTGTC.....TTGATGGTTTTTTTTTT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 264399 seqs, 353985056 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: embl-new11

1: BCT 2: FUN 3: INV1 4: INV2 5: INV3 6: MAM 7: ORG 8: PLN

9: PRI1 10: PRI2 11: PRI3 12: PRO1 13: PRO2 14: ROD 15: SYN

16: UNC 17: VRT 18: VIR

Database: genbank91

19: BCT1 20: BCT2 21: BCT3 22: BCT4 23: BCT5 24: BCT6 25: BCT7

26: INV1 27: INV2 28: INV3 29: INV4 30: INV5 31: MAM1 32: MAM2

33: PAT1 34: PAT2 35: PAT3 36: PHG 37: PLN1 38: PLN2 39: PLN3

40: PLN4 41: PLN5 42: PLN6 43: PLN7 44: PRI1 45: PRI2 46: PRI3

47: PRI4 48: PRI5 49: PRI6 50: PRI7 51: PRI8 52: PRI9 53: ROD1

54: ROD2 55: ROD3 56: ROD4 57: ROD5 58: ROD6 59: ROD7 60: STR

61: SYN 62: UNA 63: VRL1 64: VRL2 65: VRL3 66: VRL4 67: VRL5

68: VRL6 69: VRT1 70: VRT2 71: VRT3

Database: genbank-new11

72: BCT1 73: BCT2 74: INV1 75: INV2 76: MAM 77: PHG 78: PLN

79: PRI1 80: PRI2 81: PRI3 82: ROD 83: STR 84: SYN 85: UNA

Database: u-emb144\_91

88: part1

Statistics: Mean 12.445; Variance 5.270; scale 2.361

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description	Pred. No.
1	2619	94.5	2863 47	HSU00803	2863 47	HSU00803	2863 47	Human SRC-like tyrosi	0.00e+00
2	2461	88.8	2810 47	HSU22322	2810 47	HSU22322	2810 47	Human nuclear tyrosin	0.00e+00
3	1343	48.5	4590 59	RNO09583	4590 59	RNO09583	4590 59	Rattus norvegicus Spr	0.00e+00
4	1148	41.4	2602 54	MMIYKMR	2602 54	MMIYKMR	2602 54	M. musculus mRNA for i	0.00e+00
5	1066	38.5	2028 55	MUSBSK	2028 55	MUSBSK	2028 55	Mus musculus protein	0.00e+00
6	194	7.0	1829 30	SLSRK1	1829 30	SLSRK1	1829 30	S.lacustris srfk1 mRNA	1.95e-148
7	189	6.8	1702 30	SLSRK4	1702 30	SLSRK4	1702 30	S.lacustris srfk4 mRNA	1.25e-143
8	176	6.4	1253 30	SLSRK2	1253 30	SLSRK2	1253 30	S.lacustris srfk2 mRNA	3.57e-131
9	164	5.9	1190 30	SLSRK3	1190 30	SLSRK3	1190 30	S.lacustris srfk3 mRNA	9.78e-120
10	162	5.8	182 88	RNO2888	182 88	RNO2888	182 88	Rattus norvegicus Spr	7.77e-118
11	127	4.6	1820 71	U01350	1820 71	U01350	1820 71	Torpedo californica p	6.19e-85
12	121	4.4	2019 29	HYDSYK	2019 29	HYDSYK	2019 29	H.attenuata src-relat	2.24e-79
13	117	4.2	1518 50	HUMLYNTK	1518 50	HUMLYNTK	1518 50	Human Lyn B protein (	1.10e-75
14	117	4.2	2298 50	HUMLYN	2298 50	HUMLYN	2298 50	Human lyn mRNA encodi	1.10e-75
15	110	4.0	4550 53	MMCYES	4550 53	MMCYES	4550 53	M.musculus c-yes mRNA	2.89e-69
16	108	3.9	2200 58	RNFGR	2200 58	RNFGR	2200 58	R.norvegicus FGR mRNA	1.94e-67
17	107	3.9	4517 49	HUMCYES1	4517 49	HUMCYES1	4517 49	Human c-yes-1 mRNA.	1.58e-66
18	105	3.8	1755 71	U01349	1755 71	U01349	1755 71	Torpedo californica p	1.04e-64
19	106	3.8	1982 70	GGYRKA	1982 70	GGYRKA	1982 70	G.gallus yrk mRNA.	1.28e-65
20	99	3.6	100 10	HS63F2F	100 10	HS63F2F	100 10	H.sapiens CpG DNA, cl	2.81e-59
21	99	3.6	100 79	HS63F2F	100 79	HS63F2F	100 79	H.sapiens CpG DNA, cl	2.81e-59
22	101	3.6	1804 52	HUMXYES2	1804 52	HUMXYES2	1804 52	human c-yes-2 gene.	4.39e-61
23	100	3.6	1852 71	XLYES	1852 71	XLYES	1852 71	Xenopus laevis mRNA f	3.52e-60
24	98	3.5	100 81	HS57H7R	100 81	HS57H7R	100 81	H.sapiens CpG DNA, cl	2.24e-58
25	98	3.5	100 10	HS57H7R	100 10	HS57H7R	100 10	H.sapiens CpG DNA, cl	2.24e-58
26	98	3.5	2431 71	XLFYNR	2431 71	XLFYNR	2431 71	Xenopus fyn mRNA enco	2.24e-58
27	98	3.5	2446 71	XELFYNC	2446 71	XELFYNC	2446 71	X.laevis c-fyn protei	2.24e-58
28	96	3.5	2526 71	XHCFYN	2526 71	XHCFYN	2526 71	Xiphophorus c-fyn (Xf	1.40e-56
29	97	3.5	3438 58	RATLYNTYRX	3438 58	RATLYNTYRX	3438 58	Rat lyn protein non-r	1.77e-57
30	93	3.4	1538 58	RATLYNBTYR	1538 58	RATLYNBTYR	1538 58	Rat lyn B protein tyr	6.83e-54
31	93	3.4	1601 58	RATLYNATYR	1601 58	RATLYNATYR	1601 58	Rat lyn A protein tyr	6.83e-54
32	93	3.4	1626 57	MUSSRCP6	1626 57	MUSSRCP6	1626 57	Mouse neuronal proto-	6.83e-54
33	93	3.4	2044 56	MUSLYN	2044 56	MUSLYN	2044 56	Mouse lyn protein mRN	6.83e-54
34	92	3.3	2133 53	MMCFGRMR	2133 53	MMCFGRMR	2133 53	Murine c-fgr mRNA.	5.34e-53
35	92	3.3	2179 53	MMCFGR	2179 53	MMCFGR	2179 53	Murine c-fgr53 proto-	5.34e-53
36	89	3.2	2435 52	HUMSLK	2435 52	HUMSLK	2435 52	Human src-like kinase	2.50e-50
37	90	3.2	2647 49	HUMCSYNA	2647 49	HUMCSYNA	2647 49	Human c-syn protoonco	3.23e-51
38	88	3.2	2703 56	MUSLYNB	2703 56	MUSLYNB	2703 56	Mouse lyn B protein t	1.93e-49
39	88	3.2	2766 56	MUSLYNA	2766 56	MUSLYNA	2766 56	Mouse lyn A protein t	1.93e-49
40	86	3.1	102 79	HS63F2R	102 79	HS63F2R	102 79	H.sapiens CpG DNA, cl	1.13e-47
41	86	3.1	102 10	HS63F2R	102 10	HS63F2R	102 10	H.sapiens CpG DNA, cl	1.13e-47
42	85	3.1	1701 70	GCYES	1701 70	GCYES	1701 70	Chicken mRNA for p61	8.63e-47
43	85	3.1	1805 70	GGYES	1805 70	GGYES	1805 70	G.gallus mRNA for yes	8.63e-47
44	86	3.1	1960 53	MMHCK	1960 53	MMHCK	1960 53	Mouse hck gene for ty	1.13e-47
45	86	3.1	2002 55	MUSBMK	2002 55	MUSBMK	2002 55	Murine macrophage gen	1.13e-47

## ALIGNMENTS

RESULT	1	HSU00803	2863 bp	mRNA	PRI	25-MAY-1994
LOCUS		Human SRC-like tyrosine kinase				
DEFINITION						
ACCESSION		U00803				
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;				
		Eutheria; Primates; Catarrhini; Hominidae; Homo.				

MKK3

REFERENCE 1 (bases 1 to 2863)  
AUTHORS Lee, J., Wang, Z., Luoh, S.M., Wood, W.I. and Scadden, D.T.  
TITLE Cloning of FRK, a novel intracellular SRC-like tyrosine kinase-encoding gene  
JOURNAL Gene 138, 247-251 (1994)  
MEDLINE 9474404  
REFERENCE 2 (bases 1 to 2863)  
AUTHORS Scadden, D.T.  
TITLE Direct Submission  
JOURNAL Submitted (17-AUG-1993) D.T. Scadden, New England Deaconess Hospital, Hematology/Oncology, 185 Pilgrim Road, Boston, MA, USA, 02215

COMMENT NCBI gi: 392887  
FEATURES Location/Qualifiers  
source 1..2863  
/organism="Homo sapiens"  
/cell\_line="BL979"  
/cell\_type="lymphocyte/hepatoma"  
/tissue\_type="lymphoid"  
448..1965  
CDS  
/gene="FRK"  
/note="NCBI gi: 392888"  
/codon\_start=1  
/product="SRC-like tyrosine kinase"  
/translation="MSNIQRLMEYLEPILSLTSEADKSTVVENPCLSPQSQRHG  
HYVALFDYQARTAEDELFRAGDKLQVLDLTHGFMFARHLEKRRDGSQQIQGYIPS  
NYVAEDRSIQAEWFFGAIERSDAEKQLLYSENKTSGLFIIRESSEKGEFSLVLDGA  
VVKHYRKLREDEGGFTLRRIFSTLINEFVSHYKTSGLCVKLGKPKIKIOPAPFD  
LSYKTVDDMEIDRNSQLIKRLAGSGQFGEWEGELWNTPVAVKTLKPGSDMPNDFLR  
EAOIMKNLRHPKLIQIYAVCTLEDPIYIIITELMEHGSLOEYLDNDGSKIHILTOQVDM  
AAQVAGWAYLESRYNIHRDLAARNVLGERNIYKVADEGLARVFKVDMEDIYESRHE  
IKLPKWTAPAEIRSNKFSIKSDVMSFGILLYEIIYTKMPYSGMTCAQVIMLAQNY  
RLPQPSNCPQQFNYIMLECNWAEKPERPTFTFLMKLEDYETDSDSYDANNFIR"

BASE COUNT 895 a 550 c 624 g 794 t  
ORIGIN  
Query Match 94.5%; Score 2619; DB 47; Length 2863;  
Best Local Similarity 99.6%; Pred. No. 0.00e+00;  
Matches 2666; Conservative 0; Mismatches 5; Indels 7; Gaps 6;

Db 164 gataagaagaagaacaccttctagtgagcagctgccagctctcgtcagtttgct 223  
Qy 82 GATAAGAAGAAAGACACCTTCTAGTGAGCAGCTGCCAGCTCTGCTCAGTTTGCT 141  
Db 224 cgggttagcaccctccagcacagaagcggtaagctctctcaggtaggacttgct 283  
Qy 142 CGGGGTAGCACTCCAGCACAGAAAGCCGGTAAGTCTCTCCAGGTAGACTTGCT 201  
Db 284 gcaaccacagctgctgactgatctgaacgggactttgcatactctccgaagtatggtga 343  
Qy 202 GCAACCCAGCTGCTGCACTGATCGAAACGGGACTTTGCATATCTCCGAAGTATGCTGA 261  
Db 344 gttagtgctgactcaaatgctgctggcgaagataaagtgagcgcagagactaag 403  
Qy 262 GTTGCTGCTGACTTCAAAAGTTCCTCTCTCCACACAGGCAATGAGTGGATCGCAGACTAAG 321  
Db 404 gggagaggagaacccctgctctcttctccccaccaggcacaatgagcaacatctgtc 463  
Qy 322 GGGAGAGGAGAGCCCTGCTCTCTCTCCACACAGGCAATGAGTGGATCGCAGACTGCTC 381  
Db 464 agaggctctgggagtagactagaacccctatctccccctgtttgtccacggaggcagcaagt 523  
Qy 382 AGAGGCTCTGGGAGTAGTACAGACCCCTATCTCCCCCTGTTTGTGCCAGGGGACAGCAAGT 441

Db 524 caacgtgattgaaatccaggggccctttgctctccccagtcacagaggcatggccact 583  
Qy 442 CAACCGTGATTGAAATCCAGGGGCCCTTTGCTCTCCCCAGTCACAGAGCAGTGGCCACT 501  
Db 584 actttgtgcttggttgattaccaggctcgagctgctgaggacttgagcttcgcgagcag 643  
Qy 502 ACTTGTGTGGCTTTGTTGATTACGAGGCTCGGACTGCTGAGGACTTGAGCTTCGAGCAG 561  
Db 644 gtgacaacttcaagttctggacacttgcagggcgtggtggtttgtccagacacttgg 703  
Qy 562 GTGCAAACTTCAAGTCTGAGCACTTTGCATGAGGGCTGGTGGTTTCCAGACACTTGG 621  
Db 704 agaaaagcagagatgctccagtccagcaactacaaggctatatctctcttaactacgtgg 763  
Qy 622 AGAAAAGCAGCAGATGGCTCCAGTCAGCAACTACAAGGCTATATTCCTTCTAACTACGTGG 681  
Db 764 ctgaggagcagaagcctcaggcagagccgctggtctttggagcaatcggaaagatcagatg 823  
Qy 682 CTGAGGACAGAAAGCCTACAGGCAGAGCCGTGCTTCTTGGAGCAATCGGAAGATCAGATG 741  
Db 824 cagagaacaactatataattcagaacaagaccggttcccttcttaataatcagagaaagt 883  
Qy 742 CAGAGAAACAATATATATTCAGAAACAGACCGGTTCTTTTCTAATCAGAGAAAGTG 801  
Db 884 aaagccaaaaagagaattctctcttccagttttagatggagcagttgttaaaactaca 943  
Qy 802 AAAGCCAAAAGCGAATCTCTCTCTTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 861  
Db 944 gaattaaaagactgagtagaagggttttttctcacgcgaagaagaatcttttcaaac 1003  
Qy 862 GAATTAAGAACTGGATGAAGGGGGATTTTCTCACCGCAAGAAATCTTTTCAACAC 921  
Db 1004 tgaagaatttggagccactaccacaagaagtgacggcctgtgtcaagctgggga 1063  
Qy 922 TGAACGAATTTCTGAGCCACTACACCAAGCAAGTGACGGCCTGTGTCAAGCTGGGA 981  
Db 1064 aaccctgttaagatccaggctccagctccatttgattgtgtataaaaacgtggacc 1123  
Qy 982 AACCATCTTAAGATCCAGGTCGCCAGCTCCATTTGATTTGCTGATATAAAACCGTGGACC 1041  
Db 1124 aatggagatagaccgcaactccatcacgttctgaagcagttgggactggtcagtttg 1183  
Qy 1042 AATGGGAGATAGACCGCACTCCATACAGCTTCTGAGCCATTTGGGATCTGTCAGTTTG 1101  
Db 1184 gcgaagtatgggaaggtctgtggaaacaataaccactccagtagcagtgaacaattaaaac 1243  
Qy 1102 CGGAAGTATGGGAAGGCTCTGTGGAAACAATACCCTCCAGTAGCAGTGAAGAACTTAAAC 1161  
Db 1244 caggttcaatggatccaaatgacttccaggaggcagcagatgaagaacctaaagac 1303  
Qy 1162 CAGGTTCAATGCATCCAAATGACTTCTCAGGGAGGCAAGATAATGAAGACCTTAAGAC 1221  
Db 1304 atccaaagcttatccagctttatgctgtttgcaactttagaagatccaatttatatta 1363  
Qy 1222 ATCCAAAGCTTATCCAGCTTTATGCTGTTTGGACTTTAGAGATCCAAATTTATATTA 1281  
Db 1364 cagagttgatgagacatggaagctctgcaagaatatctccaaaatgacactggatcaaaaa 1423  
Qy 1282 CAGAGTTGATGAGCATGGAAGTCTGCAGAGATATCTCCAAAATGACACTGGATCAAAA 1341  
Db 1424 tccatctgactcaacaggttagacatggcgccagaggttgcctctggaatggcctatctgg 1483  
Qy 1342 TCCATCTGACTCAACAGGTAGACATGGCGGACAGAGTTCCTCTGGAATGGCCTATCTGG 1401

Db 1484 agtctcggaactacattcacagagatctggctgccagaaatgctcctggtggtgaacata 1543  
Qy 1402 AGTCTCGGAACACATTCACACAGATCTGGCTGCCAGAAATGCTCGTGGTGAACATA 1461  
Db 1544 atattctacaagtgcagagattttggacttgcagagatttttaagtgagataatgaagaca 1603  
Qy 1462 ATATCTACAAGTAGCAGATTTTGGACTTGCACAGATTTTAAAGGTAGATAATGAAGACA 1521  
Db 1604 tctatgaaatcagacgaataaagctgcggtggaagtgcgcgcgaagccattc 1663  
Qy 1522 TCTATGAATTAGACACGAAATPAAAGCTGCGGTGAAGTGAGCTGCCGCCGAGCCATT 1581  
Db 1664 gtatgaataaattcagcatttaagtcgcatgtatggtcatttggaaatccttcttttgaaa 1723  
Qy 1582 GTAGTAATPAAATCAGCATTAAGTCCGATGTATGTCATTTGGAAATCCTTCCTTTATGAAA 1641  
Db 1724 tcattacttggcaaatgcttcacagtggtatgcaggtgcccagtgtaattccagatgt 1783  
Qy 1642 TCATTACTTATGGCAAAATGCCCTTACAGTGGTATGACAGGTGCCAGGTGAATCCAGATGT 1701  
Db 1784 tggctcaaaactatagacttcgcaaccatccaactgtccacagcaattttacaacatca 1843  
Qy 1702 TGGCTCAAAACTATAGACTTCCGCAACCATCCAACTGTCACAGCAATTTTACAACATCA 1761  
Db 1844 tgttgagtgctggaatgcagagcttaaggaacgacctacatttgagacac'tgcgttggga 1903  
Qy 1762 TGTGGAGTGCTGGAATGCAGAGCCTTAAGGAAGCAGCTACATTTGACAGACTGCCGTGGA 1821  
Db 1904 aacttgaagactattttgaacagactcttcacattcagatgcacaaataaactcataagat 1963  
Qy 1822 AACTTGACAGACTATTTTGAACAGACTCTTCATATTCAGATCGAATAAATCACTTCATAGAT 1881  
Db 1964 gaacactggagaagatatcaaatataaagttagcaaaaacaattcaaatccattcc 2023  
Qy 1882 GAACACTGGAGANGAATATCAAAATAAAGTAGCAAAAACAAATTCAAATATCCATTCC 1941  
Db 2024 aaatacaatgtatacaaacactgcacatcagtttactcctgacatattcaagtgtag 2083  
Qy 1942 AAAATACAATGTATACAACCACTGCACAACTCAGTTATCTGCATATTTCAAGTGATAG 2001  
Db 2084 gataaagttggccatgtattatgaaaagattattgtgcattttattgactgggcaaca 2143  
Qy 2002 GATAAAGTTGGCCATGTATTATGAAAAAGATTATTGTGCAATTTATTGACTGGCACA 2061  
Db 2144 ctgcaggacagtcgaagtgatataatttccctcaactgcctggtgtaaaatgaagcaccta 2203  
Qy 2062 CTGCAGCAGACTGAAGCTCATATATATTCG-TCAGTGCCTGG-AAAATTAAGCAGACTA 2119  
Db 2204 aaccaagttattttcttttaagagataacttacatttccatttattgtttgaaatdgc 2263  
Qy 2120 AACCAAGTATTTTCTTTTAAAGAGATACTTACATTTCCATTATTGTTGAAATGTGC 2179  
Db 2264 ---atcaagaatacaacagatgagtccaatttttactcagtgatg-t-tgtagcatt 2319  
Qy 2180 CGATCAAGGAATCAACAGATCATAGTCCAAATTTTACTCAGTCACTGCTGTAGCATT 2239  
Db 2320 ttcctgttactgattagagtggttattcatttctcagatttgcgtgaatcccatcagg 2379  
Qy 2240 TTCTCTGTTTACTGATTAGAGTGGTTATTCAATTATTCCTCAGATTCTCGAATCCCATCAGG 2299  
Db 2380 ctgttattatgaaggaatttattgttcttgcgcagcaggaacctgtgctttgagattt 2439  
Qy 2300 CTGTTATTATGAGGAATTTGATTGCTTTGCTGCGACAGGAGGACTGCTGCTTTGAGATTT 2359

Db 2440 ttttttctcttttaaaatacctcgtgaactacaatgatggtaaagccatgtttaaatgactt 2499  
Qy 2360 TTTTCTCTCTTTAAAAATATCCTGTAACATACAATGATGCTAAAGCCATGTAAATGACTT 2419  
Db 2500 gattgtacttggagtaattgcacattttttctctatgcatataaaaaaatgatgcagctgttg 2559  
Qy 2420 GATTGTACTTGGGTAAATGTCACATTTTTTTCTATGTCATAAAAAAATGATGCAGCTGTG 2479  
Db 2560 agaaaacgaagtctttttcatttgcagaagaaatgatggaatttttctgtacttcagt 2619  
Qy 2480 AGAAAACGAAGTCTTTTTTCATTTTGGCAGAGAAATGATGGAATTTTTCTGTACTTCACT 2539  
Db 2620 atgtgtcaactgagagtcatacacattagtttttaactctcttaattattgagaatcaggttg 2679  
Qy 2540 ATGTGTCAACTCAGAGTCATATACATTAGTTTAAATCTCTTAATATTGAGAATCAGGTG 2599  
Db 2680 caaaacgagatgagttatttctatggaatgtgagaaatgtctaatagcccatataaagtct 2739  
Qy 2600 CAAAACGGAATGAGTTATTATCTATGGAATGTGCAAAATGCTAATAGCCCCATAAAGTCT 2659  
Db 2740 gagaataatgatatcaaaatagtttagaagaaatgagagagagacagtag-attgctgtggc 2798  
Qy 2660 GAGAAATAGGTATCAAAATAGTTTATAGAAAATGACAGACAGACAGTAGGATTCTGTGCG 2719  
Db 2799 ctgaacttctgagtaattaataaagaaaaagaagtacc 2836  
Qy 2720 CTAGACTCTGAGTAATTAATAAGAAAAGAGATACC 2757  
RESULT 2  
LOCUS HSU22322 2810 bp mRNA PRI 27-MAR-1995  
DEFINITION Human nuclear tyrosine protein kinase Rak mRNA, complete cds.  
ACCESSION U22322  
KEYWORDS .  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotea; mitochondrial eukaryotes; eukaryote crown group;  
Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2810)  
AUTHORS Cance,W.G., Craven,R.J., Bergman,M., Xu,L., Alitalo,K. and Liu,E.T.  
TITLE Rak, a novel nuclear tyrosine kinase expressed in epithelial cells  
JOURNAL Cell Growth Differ. 5 (12), 1347-1355 (1994)  
MEDLINE 95210168  
REFERENCE 2 (sites)  
AUTHORS Cance,W.G., Craven,R.J., Weiner,T.M. and Liu,E.T.  
TITLE Novel protein kinases expressed in human breast cancer  
JOURNAL Int. J. Cancer 54 (4), 571-577 (1993)  
MEDLINE 93293373  
REFERENCE 3 (bases 1 to 2810)  
AUTHORS Cance,W.G.  
TITLE Direct Submission  
JOURNAL Submitted (06-MAR-1995) William G. Cance, Surgery Dept., University  
of North Carolina, Chapel Hill, NC 27599, USA  
COMMENT NCBI gi: 732527  
FEATURES Location/Qualifiers  
source 1..2810  
/organism="Homo sapiens"  
/cell\_line="BT-20 breast cancer cell line"  
410..1927  
CDS  
/note="nuclear tyrosine protein kinase; Method:  
conceptual translation supplied by author. NCBI gi:





QY 1676 GACAGTGCAGGTAATCCAGATGTTGGCTCAAACTATAGACTTCGCCACCATCCAA 1735  
Db 1780 ctgtccacagcaattttacacatcatgttgagtgctggaatgcagagcctcaaggaacg 1839  
QY 1736 CTGTCACAGCAATTTTACAAATCATGTGTGGAGTCTGGAATGCCAGAGCTTAAGAACG 1795  
Db 1840 acctacatttgagacactcggttggaacttgagactatttttgaacagactcttcata 1899  
QY 1796 ACCTACATTTGAGACACTGCTGTGAACTTGAGACTATTTTGAACACAGACTCTTCATA 1855  
Db 1900 ttcagatgcaaatcaactcaatgaatgaacactggagagaatatcaaatcaataaagtag 1959  
QY 1856 TFCAGATGCAATAAATTCATAGATGAACACTGCGAGAGAAATATCAATAATAAGTAG 1915  
Db 1960 caaacaataatcaataataatccattccaaaatacaaatgttatcaaccaactgcacaa 2019  
QY 1916 CAACAATAATCAATAAT---CAATTCCAATAATACATGTTATCAACCAACTGCACAA 1972  
Db 2020 cagtttatcctgacataattcaagtgtatgataaagttggccatgtattatgaaaagat 2079  
QY 1973 CAGTTTATCTCACATATCAAGTATAGATGAATAAGTTGGCCATGATATATGAAAAAGAT 2032  
Db 2080 tatitgtcattttatigactgggcaacactgcagacagcgaagtgtatataatttc 2139  
QY 2033 TATTTGTGATTTTATTCCTGGCAACACTGCAGACAGTCAAGTCAATATAATATTCG 2092  
Db 2140 ctcaactcgtgtgaaatgaacacatacaacaaagttattttttttaaagataact 2199  
QY 2093 -TCACCTGCTGG-AAAATTAAGACACTAACCAAGTATTTTCTTTTAAAGACTACT 2150  
Db 2200 tacattccattttattgttgaatgtcg---atcaagagaatacaacagatgatgtccaa 2257  
QY 2151 TACATTTCCATTTATTTTGAATGTCGGATCAAGAAATCAACAGATGATGTCGA 2210  
Db 2258 tttttactcagtgactg-t-tgtagcattttctgttactgattagtggttat-cat 2314  
QY 2211 TTTTACTCAGTGAAGTGTGTAGCAATTTCTGTTACTGATTAAGATGGTTATCAT 2270  
Db 2315 tattcctcagattgctgaatcccatcagctgttattatgaaggaatttgattgtctgc 2374  
QY 2271 TATTCCTCAGATTGCTGAATCCCATCAGGCTGTATTATCAAGGAATTCATGCTTTCG 2330  
Db 2375 tgcacagcaggacctgtgctttgagattttttttctcttttaaaatcctgtgaactac 2434  
QY 2331 TGCACAGCAGGACCTGTGCTTTGAGATTTTTCCTTTTAAAATATCCTGTAAGTAC 2390  
Db 2435 aatgatggttaagcactgttaaatgacttgattgtacttgagtaattgcacatttttt 2494  
QY 2391 AATGATGTTAAGCAATGTTAAATGACTGATGTTACTTGGAGTAATGTCACATTTTT 2450  
Db 2495 tctatgcataaaaaaatgacagctgtgagaaaacgaagcttttttctatttgcagaa 2554  
QY 2451 -CTATGCTAATAAAAAATGATGCGCTGTTGAGAAAACGAAGTCTTTTCAATTTTCGAAA 2509  
Db 2555 ggaatgatggaatttttctgtactcagatgtgtcaactgagagtcataacatagt 2614  
QY 2510 GAAAATCATCGAATTTTCTGTACTTCACTATGCTCAACTGACATCATACATAGT 2569  
Db 2615 ttttaactcttaatttgagaactcaggttgcaaa-cggatgagttattatctatg-aaat 2672  
QY 2570 TTTTAATCTCTAAATTTGAGAAATCAGGTTGCAAAACGATGATTTATATCTATCGAAT 2629  
Db 2673 gtgagaaatgtctaagcccaataa-gtctgagaaataggtatcaaaaatagtttaggaaa 2731

QY 2630 GTGAGAAATGCTAATATAGCCCATAAAGTCTGAGAAATAGGTATCAAAAATAGTTTAGGAAA 2689  
Db 2732 atgagagagaacagtaggattgctgtggcctagacttcg-gtaatttaataaagaaaa 2790  
QY 2690 ATCAGAGAGAACAGTAGGATTTGCTGTGGCTAGACTTCTGAGTAATTAATAAGAAAAA 2749  
Db 2791 gaagtacc 2798  
QY 2750 GAAGTACC 2757  
RESULT 3  
LOCUS RNU09583 4590 bp mRNA ROD 09-AUG-1995  
DEFINITION Rattus norvegicus Sprague-Dawley src related tyrosine kinase mRNA,  
complete cds.  
ACCESSION U09583 U02888  
KEYWORDS .  
SOURCE rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Eukaryotae; Hyperchondria; Eukaryote crown group; Metazoa/Eumycota  
group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia;  
Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii;  
Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires;  
Rodentia; Sciurognathi; Myomorpha; Muridae; Rattus.  
REFERENCE 1 (bases 1 to 4590)  
AUTHORS Avigan,M.I. and Sunitha,I.  
TITLE Characterization of GASK, a novel src related tyrosine kinase  
JOURNAL Unpublished  
REFERENCE 2 (bases 1510 to 1691)  
AUTHORS Sunitha,I. and Avigan,M.I.  
TITLE A newly identified tyrosine kinase is preferentially expressed in  
the gastrointestinal tract  
JOURNAL Biochim. Biophys. Acta (1994) In press  
REFERENCE 3 (bases 1 to 4590)  
AUTHORS Avigan,M.I.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAY-1994) Mark M. Avigan, Pathology, Georgetown  
University School of Medicine, 3900 Reservoir Road N.W.,  
Washington, D.C. 20007, USA  
COMMENT NCBI gi: 939624  
FEATURES Location/Qualifiers  
source  
1..4590  
/clone="GASK"  
/clone\_lib="RintTKBS"  
/strain="Sprague-Dawley"  
/organism="Rattus Norvegicus"  
/sex="male"  
/tissue\_type="small intestine"  
/dev\_stage="infant"  
445..1965  
CDS  
/note="NCBI gi: 939625"  
/codon\_start=1  
/product="src related tyrosine kinase"  
/translation="MGVCMRLWAYIQLPFCWSQADKSVVIENPCAYCPPEANRSQ  
GQYFVALFDYARTAEADLSFHAGKQLVDTSHEGHWLARHLERKGGGLGQQLQYIP  
SNYVADRSIQAEPEFFGAIKRADAERQLLYSENQCAF LIRESESGDFSLVLDL  
GVKHYRIRRLDGGGFTLRKTFSTLNEFVNYTTTSDGLCVKLEKPCIKIQVTPF  
DLSYKTQWEIDRNSIQLKRLCSGQFGEVGEIHNTPVAVTKLKGSGMDPNDL  
REQAQNKSLRHPKLQIYAVCTLEDPIYITTELNRHGSQLEYLQNDGGSKIRLQVVD  
MAAQASGMAYLESQNYIHRDLARNVLVGEHNIYKVADFGLARVKVDNEDLYESKH  
EIKLPVKVTAPEAIRTNKESIKSDVNSFGILLIYELITYGKMPYSGMTGAQVLIHLGQN  
YRLPQSNCPDEQFYSIMMECNVEPKQRTFTLHKLEDYFEPDSSYSDTNFIN"

BASE COUNT 1370 a 937 c 999 g 1284 t









CC (F03097) was isolated from lymphocytic and megakaryocytic cell  
CC libraries. The gene can be used to produce recombinant LpK2, to  
CC identify other new pTK genes, or to design drugs, peptides or  
CC antisense constructs that modulate pTK activity.

SQ Sequence 7607 BP; 1954 A; 1851 C; 1693 G; 2109 T;

Query Match 94.5%; Score 2619; DB 16; Length 7607;  
Best Local Similarity 99.68; Pred. No. 0.00e+00;  
Matches 2666; Conservative 0; Mismatches 5; Indels 7; Gaps 6;

Db 987 ggtactcttttttttatttaataactcaagaagctcaggccacagaacac-tactgtttct 1045  
|||||  
Cp 2757 GGTACTCTTTTCTTTATTAATTAATCAAGAGCTAGGCCACAGAACTCACTACTGTCT 2698

Db 1046 cctctcatttctcctaactatttggatacctatttctcagaactttatggctattagaca 1105  
|||||  
Cp 2697 CCTCTCATTTTCTTAACATATTTTGTACCTATTTCTCAGACTTTATGGCTATTAGACA 2638

Db 1106 ttctcactttccatagataaactcatccgttttggcaacctgattctcaatattaag 1165  
|||||  
Cp 2637 TTCTCAGATTTCCATAGATAAATCACTCCGTTTGGCAACTCATTTCAATATTAG 2578

Db 1166 agattaaaaactatgatctcagttgacacatactgaagtacagaaaaattcca 1225  
|||||  
Cp 2577 AGATTAAACTAATGTATATGACTCTCAGTTGACACATATCAAGTACAGAAAATTTCCA 2518

Db 1226 tcatttctctcgaaatgaagaagactcgttttctcaacagctgcattcttttta 1285  
|||||  
Cp 2517 TCATTTCTCTCGAAAATGAAAAGACTTCGTTTTTCTCAACAGCTGCATCTTTTTTA 2458

Db 1286 tgcataagaaaaatgtgaattactccaagtacaatcaagtcatttaacatggctttac 1345  
|||||  
Cp 2457 TGCATAGAAAATAATGTGCAATTAATCTCAAGTACAAATCAAGTCATTAAATGGCTTTAC 2398

Db 1346 catcattgtagttacagatatttttaaaagagaaaaaaatctcaagcacaggtctcg 1405  
|||||  
Cp 2397 CATCATTTGTAGTACAGATATTTTAAAGACAAAATAATCTCAAGCAAGTCTCG 2338

Db 1406 ctgtgcagcaagaactcaaaatccttcaataaacagctgattgaggtcagcaatctg 1465  
|||||  
Cp 2337 CTGTGACGAAAGCAATCAAAATCTTCAATAATACAGGCTGATGGGATTCAGCAATCTG 2278

Db 1466 aggaataatgaataaccactctaatcagtaaacaggaataatgctaca-acag-tcactga 1523  
|||||  
Cp 2277 AGGAATAATGAATAACCACTCTAATCAGTAACAGGAAATGCTACAGCTCATCTGA 2218

Db 1524 gtaaaaattggactatcatctgtgtattctcttggatcg--acatttcaacaataaatgg 1581  
|||||  
Cp 2217 GTAAAAATTTGGACTATCATCTGTTGATTCCTGTGATCGGACATTTCAAAACAATAAATGG 2158

Db 1582 aaatgtaagtatctcttaaaagaaaaataacttggtttagtggcttaattttaccagg 1641  
|||||  
Cp 2157 AAATGTAAGTATCTCTTAAAGAAAATAAATCTGGTTTGTAGTGGCTTAATTTT-CAAGG 2099

Db 1642 cagtgaggaataatataatcaccttgactgctcagtggttgcccaagtcacaataaagtc 1701  
|||||  
Cp 2098 CAGTGAAGCAA-TTATATATGACCTTGACTGTCTCGAGTGTGGCCAGTCAATAAATAATGC 2040

Db 1702 acaataaatctttttcataatcatgccaaactttatcctatcatcacttgaatattgtcagga 1761  
|||||  
Cp 2039 ACAAAATAATCTTTTCAATAATCATGCCCCAACTTTATCTATCATCTTGAATATGTGAGGA 1980

Db 1762 taactgattgtgcagttgggtgataaacattgtattttgggaatggattattgtgaattgt 1821  
|||||

Cp 1979 TAAACTGATTTGCGAGTTGGTTGATAAATGTTATTTTGGAAATGGATTATTTGAAATTTCT 1970  
Db 1822 ttgtactcttattattgatattctctccaggttctcatcttataagattatttgcac 1881  
|||||

Cp 1919 TTTGCTACTTTATTAATTTGATATTTCTTCCAGTGTTCATCTTATGAAGTATTATTGCATC 1860  
Db 1882 tgaatagaagagctctgtttccaaaatagcttccaagtttccaacgacagctgtctcaaatgt 1941  
|||||

Cp 1859 TCAATATATGAAGAGTCTGTCTTCAAAAATAGTCTTCAAGTTTCCAACGCGCTGTCTCAAAATCT 1800  
Db 1942 agtgcgttctttaggtctgcatctccagcactccaaacatgatgtgttaaaattgctgtgg 2001  
|||||

Cp 1799 AGGTCTGTTCTTAGCTCTGCATTCCAGCACTCCAAACATGATGTTGTAAATTTGCTGTGG 1740  
Db 2002 acagtgtgattggttgcggaagctctatagttttgagccaacatctggattacctggggcacc 2061  
|||||

Cp 1739 ACAGTTGCGATGTTCCGGAAGTCTATAGTTTTCAGCCACATCTGGATTACCTGGCGACC 1680  
Db 2062 tgtcataccactgtaaggtcattttgcccataagtaagtatttcataaagaaggattccaaa 2121  
|||||

Cp 1679 TGTCTATACACTGTATAGGCCATTTTGCCTAAGTAATGATTTGATTAAGAGGATTCCAAA 1620  
Db 2122 tgaccatacatcggaacttaactgaatttattactacgaatggcttcggggcgacgtcca 2181  
|||||

Cp 1619 TCACCATATACCGCACTTAATGCTGAATTTATTTACTACGAATGGCTTCGGGGCGCACTCCA 1560  
Db 2182 cttcacggcgacgtttatttctgtctagatctcatagatgtcttcattatctacetttaa 2241  
|||||

Cp 1559 CTTCCGCGCAGCTTTATTTTCTGCTGTAGATTCTAGATGCTTTCATTATCTACCTTTAA 1500  
Db 2242 aaactctggcaagtcacaaatctgtacttcttagatatattgttcaccaacagggacatt 2301  
|||||

Cp 1499 AACTCTGCGCAAGTCBAAATCTGCTACTTTGTAGATATATGTTTCCACACAGGACATTT 1440  
Db 2302 tctggcagcagcatctctgtgaatgtattccagagatccagataggccattccagagcc 2361  
|||||

Cp 1439 TCTGGCAGGCAGATCTCTGTGAATGTAGTTCCGACATCCAGATAGGCCAATTCAGAGGC 1380  
Db 2362 aaactctgcccacatctctacactctgaagtcagatggatttttggatccagctgtcatttg 2421  
|||||

Cp 1379 AAACCTGTGCCGCCATGTCTACCTGTGAGTCAGATGGATTTTGTATCCAGTCACTATTTTG 1320  
Db 2422 gagatattcttgcagacttccatgtctcatcaactctgtaataataataatggatcttc 2481  
|||||

Cp 1319 GAGATATTCTTCAGACTTCCATGCTCTCATCAACTCTGTAATAATATAAATTTGGATCTTC 1260  
Db 2482 taaagtgcaaacagcatcaaaagctggatggatcttggatgtctttaggttcttcttcatctg 2541  
|||||

Cp 1259 TAAAGTCAAAACAGCATAAAGCTGGATAGCTTTTGGATGCTTTAGGTTCTTCATTATCTG 1200  
Db 2542 tgcctccctcaggaaagtcatttggatccattgaaacctgggttttaagtgttttccactgctac 2601  
|||||

Cp 1199 TGCTTCCTCAGGAAGTCATTGGATCCATTGAACCTGGTTTAAATGTTTTCACCTGCTAC 1140  
Db 2602 tggagtggttattgttccacagacttcccatacttcccaactgaccagatcccaatcg 2661  
|||||

Cp 1139 TGGAGTGGTATTTGTTCCAGACCTTCCCATACTTCGCCCAACTGACCAGATCCCAATCG 1080  
Db 2662 ctccaagaactgtatggagtgcggtctatctcccatgtgtccacaggtttttatacagaaa 2721  
|||||

Cp 1079 CTTCAGAAAGCTGATGGAGTTCGGGTCTATCTCCCATTTGTCACGGTTTTTATACGAAA 1020  
Db 2722 atcaaatggagctgggacctggatctttaagcatggtttccccagcttggacacagggcc 2781  
|||||



|||||  
Cp 2397 CATATTGTAGTTACAGATATTTTAAACAGAAAAAAATCTCAAGACACAGTCTCTG 2338  
Db 1406 ctgtgcagcaagaacaaatccttccataataacagcctgatggattcagcaatctg 1465  
|||||  
Cp 2337 CTGTGCAGCAAGCAATCAAAATCTTCATAATAACAGCCGTATGGGATTCAGCAATCTG 2278  
Db 1466 aggaataatgaataaacaccatctaatcagtaaacagagaaaaatgctaca-acag-tcaatga 1523  
|||||  
Cp 2277 AGGAATAATGAATAACCACTCTAATCAGTAAACAGAGAAAAATGCTACACAGTCATCACTGA 2218  
Db 1524 gtaaaatggagactatcatctgttgattctcttgatcg--acatttcaaaacaataatgg 1581  
|||||  
Cp 2217 GTAAAAATGGAGCTATCATCTCTGATTCCTTGATCGCGCAATTTCAACACATAAATGG 2158  
Db 1582 aaatgaagtatctcttaaaaaagaaaaataaacttggttttagtggcttaattttaccagg 1641  
|||||  
Cp 2157 AAATGTAGTATCTCTTAAAGAGAAAAATAACTTGGTTTAGTGTCTTAATTTT-CAAGG 2099  
Db 1642 cagtgaggaataatatacaccttgactgctcctgcagtggtgccagtgcaataaaatgc 1701  
|||||  
Cp 2098 CAGTAGCAAA-PTATATATGACCTTGACTGTCTCGCAGTGTGGCCAGCTCAATAAATGC 2040  
Db 1702 acaaaataacttttttataatacatgccaactttatcctatcacttgaatatgtcagga 1761  
|||||  
Cp 2039 ACAAAATATCTTTTCAATAATACATGCCAACTTTATCTTATCACTTGAATATGTCAGGA 1980  
Db 1762 taaactgattgvcagtggttgataaacattgtatttggaaatggatttgaatttgt 1821  
|||||  
Cp 1979 TAAACTGATTTGTGCGAGTGGTTGATTAACATTTGATTTTGGATGATTAATTTCAATTTGT 1920  
Db 1822 ttgtacttattatttattgataattcttccagtggttcacatttatgaagtatttggatc 1881  
|||||  
Cp 1919 TTTGCTACTTTATTAATTGATATCTCTCGCAGTGTCACTTTATGAAGTATATTGCGATC 1860  
Db 1882 tgaatatgaagagctgtttcaaaatagttcttcaagtttccaaagcagtgctcctcaaatgt 1941  
Cp 1859 TGAATATGAAGAGCTGTGTTTCAAAATAGTCTTCAAGTTTCCAAAGCAGTGTCTCAAAATGT 1800  
Db 1942 aggtcgttccataggtcttgatctceagcaactccaaatgatgtgtgaatttgtctgg 2001  
Cp 1799 AGGTGCTTCTTAGGCTCTGCATTCCAGCACTCCAAACATGATGTTGTAAATTTGCTGTGG 1740  
Db 2002 acagttggatgggtgoggaagtctatagtttbgagccaacatctggtattacctggcacc 2061  
Cp 1739 ACAGTTGGATGGTTGGGAAGTCTATAGTTTGGAGCAACATCTGGATTACCTGGGCACC 1680  
Db 2062 tgtcataccactgtaagcgatttggccataaagtatttccataaagaaggtatccaaa 2121  
Cp 1679 TGTCTATACCACTGTAGGCAATTTGCCATAAGATTAATGATTTCTAATAAAGAGATTCCAAA 1620  
Db 2122 tgaccatcacatcgactaatgtcgaatttattactacgaatggcttcggcgagtcctca 2181  
Cp 1619 TGACCATATACCGGATTAATGCTGAATTTATTACTACGAATGGCTTCGGGCGCAGTCCA 1560  
Db 2182 ctccacggcgagtttatttctgtctagattcatagatgtctcattatctacctttaa 2241  
Cp 1559 CTTTCAACGGCAGCTTATTTCTGTCTAGATTCTATAGATTCTCTTATTAATCTACCTTAAA 1500  
Db 2242 aactctggcaggtccaaaatctgtaacttctgtagatattatgttccaccaacagggacatt 2301  
|||||  
Cp 1499 AACTCTGGCAAGTCCAAAATCTGCTACTTTGTAGATATTATGTTCCACCAACGAGCAATT 1440  
Db 2302 tctggcagcagatctctgtgaatgtagtttccagactccagatagggcatttccagaggc 2361

|||||  
Cp 1439 TCTGGAGCCAGATCTCTGTGAATGTAGTTCCGAGCTCCAGATAGGCCATTCAGAGGC 1380  
Db 2362 aaactgtgcgcgcaatgtctacctgttgaqtcagatggatttttgcacagtgctcattttg 2421  
|||||  
Cp 1379 AACCTGTGCCGCATGTCTACCTGTGAGTCAGATGGATTTTGTGATCCAGTGTCAATTTG 1320  
Db 2422 gagatattcttcagagacttccatgtctcatcaactctgtaataataaattggatcttc 2481  
|||||  
Cp 1319 GAGATATTCTTCGACACTCCCATGTCTCATCAACTCTCTATAATAATAAATTTGGATCTTC 1260  
Db 2482 taaagtcaaacagcagatcaagctggatgaagcttttggatgtcttagtttcttcaattatctg 2541  
|||||  
Cp 1259 TAAAGTGCAAAACAGCATTAAGCTTGATAGCTTTGGATGCTCTTAGGTTCTTCATTATCTG 1200  
Db 2542 tgcctccctcaggaaagtcatttggatccattgaaacctgggttttaagtgttttcaactgtac 2601  
|||||  
Cp 1199 TGCTCTCCCTCAGGAAGTCATTTGGATCCATTGAACCTGGTTTAAATGTTTTCAC7GCTAC 1140  
Db 2602 tggagtgttattgttccacagaccttcccatattcgcgaactgacacagatcccaatcg 2661  
|||||  
Cp 1139 TGGAGTGGTATTGTTTCCACAGACCTTCCATATCTTCGCCAAACTGACACAGATCCCAATCG 1080  
Db 2662 cttcagaagctgtatggagtcgggtctatctccatgtgcacaggttttatacagaaa 2721  
|||||  
Cp 1079 CTTCAGAAGCTGTATGGAGTTCGGCTCTATCTCCATTTGTCACGGTTTATACGACAA 1020  
Db 2722 atcaaatggagctgggacctggatctttaagcatggtttcccccagcttgacacagggcc 2781  
|||||  
Cp 1019 ATCAAAATGGAGCTGGGACCTGATCTTTAAGCATGCTTCCGCCAGCTTGACACACAGGCC 960  
Db 2782 gtccattgtctgggtgtagtgctcacaattcgttcagtggtgaaaagattcttcttcg 2841  
Cp 959 GTCACTTGTCTTGGTGTAGTGGCTCACAAATTCGTTCAGTGTGTAAAGAGATTCTTCTTCG 900  
Db 2842 cgtgagaaaaatcccccttccacagctctttaaattctgtagttttacaactgctcc 2901  
Cp 899 COTGAGAAAAATCCCTCTATCCAGTCTTTTAAATTCGTAGTGTTTTACAAC7GCTCC 840  
Db 2902 atctaaactgaagagagaattctccttttggcttccacttctctctgattagaaagga 2961  
Cp 839 ATCTAAAAC7GAAAAGACAGATTCCTCTTTTGGCTTTTCACTTCTCTGATTAGAAAGGA 780  
Db 2962 accggtcttgtttctgaaataataatagttgtttctgcatctgatctccgattgctcc 3021  
Cp 779 ACCGGCTTGTCTTTCGATATAATAGTTGTTCTCTGCA7GCTCTCCGATTCCTCCGATTCCTCC 720  
Db 3022 aaagaacacagctctgcctgtaggcttctgtcctcagccacagtagtttagaagaataata 3081  
Cp 719 AAAGACACAGGCTCTGCTGTAGGCTCTGTCTCAGCCACGCTAGTTAGAGGAATATA 660  
Db 3082 gctctgtattgtcgtgactgagccatctcgtcttttccaaagtgtctggcaaacacca 3141  
Cp 659 GCCTTGTAGTTGCTGACTGGAGCCATCTGCTCTTCTTCCAAAGTCTCTGGCAAAACACCA 600  
Db 3142 gccctcatgcaaatgtccagaaacttgaagtttgcactgtcctgcgaagctcaagtcctc 3201  
Cp 599 GCCTCATGCAAAAGTGTCCAGAACTTGAACTTTGTCACCTGTCTCGGAAGCTCAAGTCTCTC 540  
Db 3202 agcagtcagagcctggttaatacaaaagcccaaaagtagtggccatgcctctgtgactg 3261  
|||||  
Cp 539 AGCATCCGAGCTGGTATCAACAAAGCCACAAAGTAGTGGCCATGCCTCTGTGACTG 480  
Db 3262 ggsgagagcaaaagggccccctggattttcaatcacggttgacttctgctgcctccgtggacaa 3321

Cp 479 GGGAGACAAAGGCGCCCTGGATTTCAATCAGGCTGACTGTCTGCTCCGTCGACAA 420  
|||||  
Db 3322 acaggagagatagggttactccagagactctgacagatgttgctcatgtgcc 3381  
|||||  
Cp 419 ACAGGGGAGATAGGGGTTCTAGGTACTCCAGAGCCCTGACAGATGTGCTCATTTGTC 360  
|||||  
Db 3382 ttggtgggagaagagagcagggtctctccctctccctttagtctctggatccaactt 3441  
|||||  
Cp 359 TTGGTGGGAGAGAGAGGAGCGGCTCTCCCTCTCCCTTAGTCTGCGCATCCACCTT 300  
|||||  
Db 3442 atcttcctcaccaggcaacttgaagtgcagcacaactcaccatacttcggagagt 3501  
|||||  
Cp 299 ATCTTCTCTACAGGCACTTTGAAGTCAGCAGCCACTCACCATACTTCGGAGAGATG 240  
|||||  
Db 3502 caaagtcctcagatcagtcagcagctgggttcagcaagtccctactcctggagagac 3561  
|||||  
Cp 239 CAAAGTCCGCTTTCAGATCAGTCAGAGCTGGGTGCAGCAAGTCTTACTTGGAGAGAC 180  
|||||  
Db 3562 ttaccggtctgtctgtgctgaggtgctaccccgagcacaactgagcagagctg 3621  
|||||  
Cp 179 TTACCGGCTGCTTCTGCTGGTGGAGGTGCTACCCGAGGCAAACTGAGCAGAGCTG 120  
|||||  
Db 3622 ggcagctgctcaactaggaaggtgtctttcttcttctatc 3659  
|||||  
Cp 119 GCGAGCTGCTCACTAGGAGGTGCTTTTCTTCTTATC 82  
|||||

## RESULT 3

ID Q49749 standard; DNA; 149 BP.  
AC Q49749;  
DT 10-MAR-1994 (first entry)  
DE pTK gene LpTK2 partial sequence.  
KW pTK; protein tyrosine kinase; catalytic domain; c-kit; megakaryocyte;  
RW lymphocyte; amplification; primer; polymerase chain reaction; PCR; ss.  
OS Homo sapiens.  
PN W09315201-A.  
PD 05-AUG-1993.  
PF 22-JAN-1993; U00586.  
PR 22-JAN-1992; US-826935.  
PA (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.  
PI Avraham H, Cowley S, Groopman J, Scadden D;  
DR WPI; 93-320330/40.  
PT New protein tyrosine kinase genes and proteins encoded by genes -  
are of human mega-karyocytic origin  
PS Claim 2; Fig 3A; 60pp; English.  
CC pTK genes were identified using two sets of degenerative  
CC oligonucleotide primers: a first set which amplifies all pTK DNA  
CC segments (Q49743-44), and a second set which amplifies highly  
CC conserved sequences present in the catalytic domain of the c-kit  
CC subgroup of pTKs (Q49745-46). The pTK genes identified are described  
CC in Q49747-57 and R41897-02.  
CC The LpTKs are expressed in lymphocytic cells, as well as  
CC megakaryocytic cells. The partial and full-length LpTK2 gene  
CC sequences are given in Q49749 and Q49754 respectively. The  
CC protein sequence corresp. to Q49749 is claimed (claim 7) and  
CC stated as given in the specification, however is missing from  
CC the publication.  
SQ Sequence 149 BP; 37 A; 37 C; 24 G; 51 T;  
Query Match 4.7%; Score 131; DB 8; Length 149;  
Best Local Similarity 95.2%; Pred. No. 1.76e-60;  
Matches 138; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 5 gaattcttcggcgccatccattccagcgagctttattcgtgctctagattcataga 64  
|||||  
Cp 1581 GAATGGGTTCCGGCGCAGTCCACTTCACCGGAGCTTTATTTTCGTCTAGATTCATAGA 1522  
|||||  
Db 65 tgccttcattatctaccttaaaaaactctggcaagtcacaaactctgctacttctgtagatat 124  
|||||  
Cp 1521 TGTCTTCATTATCTACCTTAAAAAACTCTGGCAAGTCCCAAAATCTGCTACTTTGTAGATAT 1462  
|||||  
Db 125 tatgttcaccaacgagagacattcct 149  
|||||  
Cp 1461 TATGTTCCACCAACGAGGACATTTCT 1437  
|||||

## RESULT 4

ID T03092 standard; DNA; 149 BP.  
AC T03092;  
DT 14-FEB-1996 (first entry)  
DE Protein tyrosine-kinase LpTK2 DNA fragment.  
KW Protein tyrosine-kinase; pTK; LpTK2; agonist; cell growth;  
KW differentiation; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 1..60  
FT /\*tag = a  
PN W09527061-A1.  
PD 12-OCT-1995.  
PF 04-APR-1995; U04228.  
PR 04-APR-1994; US-222616.  
PA (GETH ) GENENTECH INC.  
PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;  
PI Wood WI;  
DR WPI; 95-366160/47.  
DR P-PSDB; R85931.  
DE Agonist antibodies which activate specific protein tyrosine  
PT kinase(s) - also activate chimeric proteins of kinase extracellular  
PT domain and Ig constant domain, useful for studying, and therapeutic  
PT modulation of, cell growth and differentiation  
PS Disclosure; Page 36; 125pp; English.  
CC DNA probes based on protein tyrosine-kinase (pTK) sequences were used  
CC to screen cDNA libraries to identify novel pTK genes. A LpTK2 gene  
CC fragment (T03092) was isolated from lymphocytic and megakaryocytic  
CC cell line libraries and encoded a peptide (R85931) showing homology  
CC to known pTKs. The gene fragment can be used to identify other  
CC new pTK genes, or to design drugs, peptides or antisense constructs  
CC that modulate pTK activity.  
SQ Sequence 149 BP; 37 A; 37 C; 24 G; 51 T;

Query Match 4.7%; Score 131; DB 16; Length 149;  
Best Local Similarity 95.2%; Pred. No. 1.76e-60;  
Matches 138; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 5 gaattcttcggcgccatccattccagcgagctttattcgtgctctagattcataga 64  
|||||  
Cp 1581 GAATGGGTTCCGGCGCAGTCCACTTCACCGGAGCTTTATTTTCGTCTAGATTCATAGA 1522  
|||||  
Db 65 tgccttcattatctaccttaaaaaactctggcaagtcacaaactctgctacttctgtagatat 124  
|||||  
Cp 1521 TGTCTTCATTATCTACCTTAAAAAACTCTGGCAAGTCCCAAAATCTGCTACTTTGTAGATAT 1462  
|||||  
Db 125 tatgttcaccaacgagagacattcct 149  
|||||  
Cp 1461 TATGTTCCACCAACGAGGACATTTCT 1437  
|||||



RESULT 5  
ID Q46688 standard; cDNA to mRNA; 1611 BP.  
AC Q46688;  
DT 23-DEC-1993 (first entry)  
DE Human pp60 c-src gene.  
KW Endothelial; tyrosine kinase protein; pp60 c-src; ss.  
OS Homo sapien.  
PN W09314193-A.  
PD 22-JUL-1993.  
PF 05-JAN-1993; US00445.  
PR 06-JAN-1992; US-820011.  
PA (UYA ) UNIV YALE.  
PI Bell L, Luthringer DJ, Madri JA, Warren SL;  
DR WPI; 93-243209/30.  
DR P-PSDB; R39705.  
PT Genetically engineered endothelial cells - which exhibit enhanced cell migration, urokinase-type plasminogen activator activity, and reduced mononuclear cell adhesion and fibronectin prodn  
PS Disclosure; Page 69-72; 91pp; English.  
CC The DNA encoding a portion or (more preferably) the entire pp60 c-src polypeptide is used to transform endothelial cells.  
CC Transformed cells produce increased amounts of pp60 c-src and have improved therapeutic properties. They migrate at faster rates than non-transformed counterparts; have an enhanced ability to inhibit the formation of thrombi and/or dissolve thrombi once they have CC formed and exhibit reduced mononuclear cell adhesion. They can also be used to improve the success of surgical procedures such as CC coronary angioplasty, heart bypass surgery, vessel graft and stent CC implantation.  
SQ Sequence 1611 BP; 334 A; 507 C; 504 G; 266 T;  
  
Query Match 3.7%; Score 102; DB 7; Length 1611;  
Best Local Similarity 64.8%; Pred. No. 7.57e-43;  
Matches 256; Conservative 0; Mismatches 136; Indels 3; Gaps 2;  
  
Db 841 ttggcaggtgtggatgggacgtggaacgtgtaccacacaggtggccatcaaaccttg 900  
Qy 1098 TTTGGCGAAGTATGGGAGGCTCTGGGAACAATACCACCTCCAGTAGCAGTGAACACATTA 1157  
  
Db 901 aagcctggcagcatgtctccagagcgtctctgcaggagcccgaggtcatgaagactg 960  
Qy 1158 AAACACAGGTTCATTCGATCCAAATGACTTCTCTGAGGAGGACACAGATATGAAGACCTA 1217  
  
Db 961 aggcagagaagctgtgcagttgtatgtgtg--gt--ttcagaggagccattacatc 1017  
Qy 1218 AGACATCCAAAGCTTATCAGCTTTATGCTGTTCGACTTTAGAGATCCATTTTATTT 1277  
  
Db 1018 gtacagagatcatgaagagggtttgtgactttctcaaggggagacaggaag 1077  
Qy 1278 ATTACAGAGTTGATGACATGCGAAGTCTGCAAGATATCTCCAAATGACATGGGATCA 1337  
  
Db 1078 tacctgcggtgcctcagctgtggacatggctctcagatcgctcagcagcatggcgtac 1137  
Qy 1338 AAAATCCATCTEACTCAACAGGTAGACATGGCGGCACAGGTTGCTCTGGAATGGCCTAT 1397  
  
Db 1138 gtggcaggtgatgaactcagctccaccgggacctctgtgcagccacacatcctcgtgggagag 1197  
Qy 1398 CTGGAGTCTCGGAACATACATTCACAGAGATCTGGCTGCCAGAAATGTCTCTGTTGGTGA 1457  
  
Db 1198 aacctggtgtgcaagtggcagactttggctggc 1232  
Qy 1458 CATATATCTACAAAGTAGCAGATTTTGGACTTCC 1492

RESULT 6  
ID Q10572 standard; DNA; 1047 BP.  
AC Q10572;  
DT 09-APR-1991 (first entry)  
DE Human Natriuretic Peptide Receptor B.  
KW NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase; hyperaldosteronism; glaucoma; guanyl cyclase.  
OS Homo sapiens.  
FR Key Location/Qualifiers  
FT Peptide 1..22  
FT /label= signal sequence  
FT Protein 12  
FT /label= mature NPRB  
FT Domain 23..455  
FT /label= extracellular domain  
FT /note= "binds natriuretic peptides A, B and C]"  
FT Domain 456..456  
FT /label= transmembrane domain  
FT Domain 479..1047  
FT /label= cytoplasmic domain  
FT /note= "GC and protein kinase activity"  
FT Modified -site 24..26  
FT /label= N-glycos site  
FT Modified -site 35..37  
FT /label= N-glycos site  
FT Modified -site 161..163  
FT /label= N-glycos site  
FT Modified -site 195..197  
FT /label= N-glycos site  
FT Modified -site 244..246  
FT /label= N-glycos site  
FT Modified -site 277..279  
FT /label= N-glycos site  
FT Modified -site 349..351  
FT /label= N-glycos site  
FT Modified -site 500..602  
FT /label= N-glycos site  
PN W09100292-A.  
PD 10-JAN-1991.  
PF 22-JUN-1990; U03586.  
PR 23-JUN-1989; US-370673.  
PA (GETH ) GENENTECH INC.  
PI Chang M, Goeddel D, Lowe D;  
DR WPI; 91-036711/05.  
DR N-PSDB; Q10324.  
PT Natriuretic protein receptor B - for diagnosis and treatment of PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.  
PS Claim 3; Fig 1; 49pp; English.  
CC The sequence was derived from the DNA encoding natriuretic peptide CC receptor B, NPRB, having guanyl cyclase (GC) activity and protein CC kinase activity. The DNA can be inserted into expression vectors CC for the prodn. of the protein, opt. after being mutated to produce CC NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr= CC 114,952). The protein (or variants) can be used in treatment of CC natriuretic peptide disorders, and also to isolate peptides using CC affinity chromatography. Antibodies with affinity for NPRB can CC also be prepd.  
SQ Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;  
  
Query Match 3.4%; Score 94; DB 2; Length 1047;  
Best Local Similarity 8.8%; Pred. No. 4.63e-38;  
Matches 84; Conservative 267; Mismatches 594; Indels 7; Gaps 7;  
  
Db 84 avdknkyhdnnngqvcvynaasvarnashwrnnnnntagavasnsakndhnytrvtg 143



Result	No.	Score	Query		ID	Description	Pred. No.
			Match	Length			
1	2059	41.2	659	11	S28912	protein-tyrosine kin	0.00e+00
2	2031	40.6	660	11	JN0471	protein-tyrosine kin	0.00e+00
3	2016	40.3	659	11	B45184	B cell progenitor ki	0.00e+00
4	1978	39.6	630	11	JU0228	protein tyrosine kin	0.00e+00
5	1791	35.8	620	11	S33253	protein-tyrosine kin	2.24e-299
6	1762	35.3	527	13	S13763	protein-tyrosine kin	6.06e-294
7	1761	35.2	619	11	A47333	T-cell-specific tyro	9.33e-294
8	1747	35.0	608	11	JU0027	protein-tyrosine kin	3.90e-291
9	1745	34.9	527	11	A55631	protein-tyrosine kin	9.23e-291
10	1736	34.7	619	11	JN0472	protein-tyrosine kin	4.47e-289
11	1719	34.4	625	11	A43030	protein-tyrosine kin	6.79e-286
12	1686	33.7	602	11	JU0215	tyrosine kinase, tec	1.01e-279

---

WAK 2





\*\*\*\*\*

and is derived by analysis of the total score distribution.

## ALIGNMENTS

RESULT	1
LOCUS	HSBMXGENE      2456 bp      RNA      PRI      22-AUG-1995
DEFINITION	H.sapiens Bmx mRNA for cytoplasmic tyrosine kinase.

**KEYWORDS** cytoplasmic: Tyrosine kinase.

**SOURCE** human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Osteichthyes; Sarcopterygii; Mammalia; Eutheria;



Qy 1272 CAAGTCCCGGACTCTGTGCTCGGGAATGGAATCTGGGAACCTGAAAGAGAGAGAT 1331  
Db 1284 tacctgttgaaggagctgggaagtggccagtttggagtggtccagctgggcaagtggaa 1343  
Qy 1332 TACCTTGTGAAGGAGCTGGGAAGTGGCCAGTTTGGAGTGGTCCAGCTGGGCAAGTGGAA 1391  
Db 1344 ggggcagdtatgatgttctgtgttaagatgatcaaggagggtccatgtcagaagaatgaatt 1403  
Qy 1392 GGGGCAGTATGATGTGCTGTTAAGATGATCAAGGAGGGCTCCATGTCAGAGATGAATT 1451  
Db 1404 ctttcaggaggcccaagactatgatgaactcagccatcccaagctgggttaatttatgg 1463  
Qy 1452 CTTTCAGGAGGCCAGACTATGATGAACCTCAGCCATCCCAAGCTGGTAAATCTATGG 1511  
Db 1464 agtgtgtcaaggagtaaccatatacatagtgactgaatataataagcaatggctgctt 1523  
Qy 1512 AGTGTGTTCAAGAGNATACCCCATATACATAGTGAATATATGAAGCAATGGCTGCTT 1571  
Db 1524 gctgaattacctgaggagtcacggaagaagacttgaaccttccagctcttaagaatgtg 1583  
Qy 1572 GCTGAATTACCTGAGGAGTCACGGAAGGACTTGAACCTTCCAGCTCTTAGAAATGTG 1631  
Db 1584 ctacagatgtctgtgaaggatggctcttggagagtcaccaattcatacaccggactt 1643  
Qy 1632 CTACGATGTCTGAAGGATGGCTTCTTGGAGAGTCAACCAATTCATACCCGGGACTT 1691  
Db 1644 ggcgtgctgaactgcttggagacagagatctgtgtgaaagtatctgaactttggaat 1703  
Qy 1692 GCCTGCTGCTAATGCTGTGTCGACAGATCTCTGTGTAAGTATCTGACTTTGGAT 1751  
Db 1704 gacaagtgatgttctgtgacagatgtcagttcagtcaggacgaagaagttccagtcac 1763  
Qy 1752 GACAAGGATGTCTGTATGACACAGATGTGCTAGTTCAGTCGGGAAGAAGTTCCAGTCAA 1811  
Db 1764 gtggtcagctccagagtggtttcattacttcaataacagcagcaagtcagactgagtc 1823  
Qy 1812 GTGCTCAGCTCCAGAGTGTTCATTACTTCAATACAGCAGCAAGTCCAGATGCGC 1871  
Db 1824 atttggatcctgatgtggagagtggtcagcctggggaagcagccctatgaactgtatga 1883  
Qy 1872 ATTTGGATCCTGATGCGGAGGTCTTACGCTGGGAGCAGCCCTATGATCTGTATGA 1931  
Db 1884 caactcccaggtggtctgaaggtctccagggccacaggtttaccggcccccacttgc 1943  
Qy 1932 CAACTCCAGGTGCTGTAAGGCTCCAGGGCCACAGGCTTTACGGCCCCCACCCTGC 1991  
Db 1944 atcggacacactaccagatcatgtacagctgctggcagcagcttccagaaaagcgtcc 2003  
Qy 1992 ATCGGACACATCTACACAGATCATGTACAGCTGCTGGCAGAGCTTCCAGAAAAGCTCC 2051  
Db 2004 cacatttcagcaactcctgtcttccattgaaccacttcgggaagaagcagcattgaag 2063  
Qy 2052 CACATTTACGCAACTCCTGCTTCCATTGAACCTTCGGGAAAAAGCAAGCATTTGAAG 2111  
Db 2064 aagaattaggagtgctgataagaatgaatatagatgctggcagcatttccattcattt 2123  
Qy 2112 AAGAAATTAGGAGTGTGATAAGATGAATATAGATGCTGGCCAGCATTTTTCATTCA 2171  
Db 2124 taaggaaagtaggaagcagaatgaatttagtagtattttaaattagttctctgtattg 2183  
Qy 2172 TAAAGAAAGTACAGGATATGTAATTTAGCTAGTATTTTAAATAGTGTCTCTGATG 2231  
Db 2184 tctattattagaatgaacaaggcaggaaaacaaagattcccttgaatttagatcaaa 2243

Qy 2232 TCTATATTTAGAAATGACAAGGAGGAGAAACAAAAGATTCCCTTGAATTTAGGTCAA 2291  
Db 2244 ttaagttaattttttttatgctgctcctgatataaacaactttccagcctatagcagaagcac 2303  
Qy 2292 TTAGTAATTTTGTGTT-ATGCTGCCCTCATATAACACTTTCCAGCCTATAGCAAGCAC 2350  
Db 2304 atttccagactgcaatagatagagactgtgttcattgtgtaagagactggcagaactgaaaa 2363  
Qy 2351 ATTTTCAGACTGCAATATAGAGACTGTCTTCATGCTGTAAGACTGAGCAGAACTGAAAA 2410  
Db 2364 ttacttatggatattcattctttttttatttatattgtcattgtcacacaataataatcac 2423  
Qy 2411 TACTTATTCGATATTCATTTCTTTTATATTGTCATTGTCACACAATTAATATATAC 2470  
Db 2424 taccaggtacagaaatgtggaaaaaataa 2453  
Qy 2471 TACCAGGTACAAAAAATAAAAAAATAAAAAA 2500  
RESULT 2  
LOCUS MUSBTCKDNA 2468 bp mRNA ROD 18-MAY-1995  
DEFINITION Mouse Bruton agammaglobulinemia tyrosine kinase (Btk) mRNA,  
complete cds.  
ACCESSION L29788  
KEYWORDS Bruton agammaglobulinemia tyrosine kinase.  
SOURCE Mus musculus cDNA to mRNA.  
ORGANISM Mus musculus  
Eukaryotes; mitochondrial eukaryotes; Metazoa/Eumycota group;  
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;  
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygia; Chonata;  
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires; Rodentia;  
Sciurognathi; Myomorpha; Muridae; Murinae; Mus.  
REFERENCE 1 (Bases 1 to 2468)  
AUTHORS Sideras,P., Muller,S., Shiels,H., Jin,H., Khan,W.N., Nilsson,L.,  
Parkinson,E., Thomas,J.D., Branden,L., Larsson,I., Paul,W.E.,  
Rosen,F.S., Alt,F.W., Vetrie,D., Smith,C.I.E. and Xanthopoulos,K.G.  
Genomic organization of mouse and human Bruton's agammaglobulinemia  
tyrosine kinase (Btk) loci  
J. Immunol. 153 (12), 5607-5617 (1994)  
COMMENT 95081608  
NCBI gi: 625143  
FEATURES  
source location/Qualifiers  
1..2468  
/organism="Mus musculus"  
/gene="Btk"  
/map="X"  
/note="NCBI gi: 625144"  
/codon\_start=1  
/product="Bruton agammaglobulinemia tyrosine kinase"  
/translation="MAAVILESIFLKSQKKTSPLNFKRILFLTVHKLSTYEYDF  
ERGRSGKGSIDVEKITCVETVPEKNPPEROIPRRGESSEMEQISIIERFPYVF  
QVYDEGPLYVPSFTEELRRKWIHLKVNIRYNSDLVKQYHPCFWDQYLCCSTAK  
NAMCQITLENRNGSLKPGSHRKTQKLPPTPEEDQILKKPLPPEPTAAPITTELRK  
VVALYDMPMANDLQLRKGEYFLEESNLPWRAKRDKNQGEVIPSNTVTEADSI  
EMWYSKMHTRSQAEQLKQEGGFIIVRDSKAGKYTVSVFAKSTGEPQVIRHY  
VVCSTPQSYLYAEKHLFTSTPELNYHNSAGLISRLKIPVSKONKAPSTAGLY  
GWEIDPKDLTFLKELGTGFGVKYQKRGYDVAIKMIRGSMSEDEFIEAKVM  
NLSHEKLVQVCTKORPFIITFYWANGCLLYLRMHRFRQQLLEKCKDCEA  
MEYLESQFLHRDLAARNCLVNDQGVKVSDFLSRYVLDDYTSVSGSKFPVRWSP  
EVLWYKFSKSDIWAFGVLMWEIYSLGKWPYERTTNSETAELHIAQGLRLYRPLASE

BASE COUNT	713 a	559 c	602 g	594 t
ORIGIN	RVTYIMYSCHWEKADERSFKILLSNILDVMDDES#			

Query Match 13.2%; Score 329; DB 55; Length 2468;  
Best Local Similarity 64.7%; Pred. No. 6.86e-243;  
Matches 723; Conservative 0; Mismatches 394; Indels 0; Gaps 0;

[illegible]

		1781	agtacaccagctctgtagcttccaaagtccagggttcacccagaagtgctta	1840
D <b>b</b>				
		1775	AGTATGTCAGTTCCAGTCGGAAACAAGTTCCAGTCAAGTGCTCAGCTCCACAGGTCTTTC	1834
Q <b>y</b>				
		1841	tgtatagcaagttcgaagcagcaaatctgcacctggtgggttttaattggtgaga	1900
D <b>b</b>				
		1835	ATTACTTCAAATACAGCACGAAGTCAGACGATGTGGCATTTGGGATCTCGATGTGGGAGG	1894
Q <b>y</b>				
		1901	tctaactcccgtgggaagatgcogtatgagagattactacaadgagacagcagaacaca	1960
D <b>b</b>				
		1895	TGTTTCAGCCTGGGAAACGAGCCCTATGACTTGTATGACAACTCCCAGGTGTTCTCGAAG	1954
Q <b>y</b>				
		1961	tgtctcaagctttaogtctctcacagcctcatctggcatcagagaggdtatataccatca	2020
D <b>b</b>				
		1955	TCTCCAGGGCCACAGGCCTTTACCGGCCCCACCTGGCATCGGACACCATTCTACCATCA	2014
Q <b>y</b>				
		2021	tgtcacagctgctgccagaagaagcagatgaacgtcc	2057
D <b>b</b>				
		2015	TGTACAGCTGCTGGCA CGA CTTCCAGAAAAGCGCTCC	2051
Q <b>y</b>				

RESULT	3
LOCUS	MUSBPK 2485 bp mRNA 17-MAR-1993
DEFINITION	Mus musculus B cell cytoplasmic tyrosine kinase (BPK) mRNA, complete cds.
ACCESSION	L08967
KEYWORDS	tyrosine kinase.
SOURCE	Mus musculus (strain C57BL/6 x DBA/2) pre-B cell cDNA to mRNA.
ORGANISM	Mus musculus
	Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE	1 (bases 1 to 2485)
AUTHORS	Tsukada, S., Safran, D.C., Rawlings, D.J., Parolini, O., Allen, R.C., Klisak, I., J., Kubagawa, H., Mohandas, T.K., Quan, S., Belmont, J.W., Cooper, M.D., Conley, M.E. and Witte, O.N.
TITLE	Deficient expression of a B-cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia
JOURNAL	Cell 72, 279-290 (1993)

COMMENT	source
NCBI gi: 192233	Location/Qualifiers
	1..2485
	/organism="Mus musculus"
	/strain="C57BL/6 x DBA/2"
	/cell_line="70 z/3"
	/dev_stage="pre-B cell"
	/sequenced_mol="cDNA to mRNA"
	137..2116
CDS	/gene="bpg"
	/note="NCBI gi: 192234"
	/codon_start=1
	/translation="bpg"

SDS

/translations="AAWVLIETIKRSQQRKSTSPINFKRFLFLVTHKUSYYEYDF  
 ERGRGSKGSDVEETIKCVETWIPKPNPVEHQIPRGCESSDMEQI1S1TERPYPF  
 QMGQGLDVSPTTEELRPWHQLKWP1RYNSDYLQYHPCPIDQCYLSCSQ7AK  
 NAMQCI1LENRNSLKPQSHRKTAKLPPTPEEQI1LKQDLPPEPTAAP1STTELRK  
 VVALYDYNPNANDQLKQKEEYF1LEESN1PMWRARDNQCEQYFIPSN1TEAEDS1  
 EYEWY5KCHTR5SOAOL1KQCKEGEGF1VDRDSKAGKYV5VFAK5TEQGOV1RIY  
 VQSTPQSOQY1LAEKH1ST1PEL1NYHQHNSAGL1SLRK1YPSVKQKNQAS1TAGLY  
 GSWED1VDQY1F1LKEIGTQFQGVVYKMGQDQYDVA1K1REGS5MEQFDEEAKVDM  
 NS1HEKVLQY1GVFC1KQRP1F1TEYKMGGLNY1LRMBR1FOTOQLLEMKDVCEA  
 MEY1LESQF1LRDL1ARNCLNVQGVKQYDFGLSR1YVLDDEYTS1SVGSKFPVWHP5P  
 EVLNT1SKFS5KSD1WAF1GV1MWEI1YS1LQKMP1YERF1TNS1ETAH1AGL1RLYR1PLASE  
 RVTY1TNS1SHKSD1SEK1LL1NS1LVNDME5"







May 31 09:15

US-08-426-509-2.rpr

3

Query Match 100.0%; Score 3727; DB 11; Length 507;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 magrgslvsrafhgdcdaeeiprvsprflrawhppvpsarmptrrwapgtqctkceht 60  
|||||  
Qy 1 MAGRGSLVSRAFHGCDSAEELPRVSPRFLRAWHPPVPSARMPTRRWAPGTQCTKCEHT 60  
|||||

Db 61 rpkpgelafrkgdvvtilaeacnksyrvkhhtsqegllaaagalrerealsadpklslm 120  
|||||  
Qy 61 RPKPGELAFRKGDDVVTILEACENKSWRVKHHTSQEGLLAAGALREREALSADPKLSLM 120  
|||||

Db 121 pwfhgklsqgeavqqlppedglflvresarhpgdyvlcvsfgrdvihyrvlhrdghlti 180  
|||||  
Qy 121 PWFHGKLSQGEAVQQLPPEDGLFLVRESARHPGDYVLCVSFGRDVIHYRVLHRDGHITI 180  
|||||

Db 181 deavffcnlmdmvehyekdkgaictklvprkkghtksaeelaraagwlinlqhltlgaq 240  
|||||  
Qy 181 DEAVFFCNLMDMVEHYSKDKGAICTKLVRPKKHGHTKSAAEELARAGWLINLQHILTLGAQ 240  
|||||

Db 241 igegefagvlqgeylgqkvavknikcdvtaqafldetavmtkmqhenlvrlilghlgaq 300  
|||||  
Qy 241 ICEGEFAGVLOGEYLGQKVAVKNIKCDVTAQAFLDETAVMTKMQHENVRLILGHILGAQ 300  
|||||

Db 301 yivmehvsknlnvflrtgralvntaqlqlfslhvaeqmeyleskklvhrdlaarnilv 360  
|||||  
Qy 301 YIVMEHVSNGNLNVLRTGRALVNTAQLQLFSLHVAEQMEYLESKKLVHRDLAARNILV 360  
|||||

Db 361 sedlvakvsdfglakaerkglslrpkwtapealkhgkftsksdwsgvllwevfsv 420  
|||||  
Qy 361 SEDLVAKVSDFGLAKAERGLDLSRLPKWTAPALKHGKFTSKSDWSGVLLWEVFSY 420  
|||||

Db 421 grapykmslkeveavekyrmeppegcpvhlmscweaeaprrppfrklaeklar 480  
|||||  
Qy 421 GRAPYKMSLKEVSEAVEKYRMEPEGPCPVHVMSSCWEAEAPRRPPFRKLAELKLAR 480  
|||||

Db 481 elrsagapasvsgqdgadgtsprsqep 507  
|||||  
Qy 481 ELRSAGAPASVSGQDAGDSTSPRSQEP 507  
|||||

RESULT 2 A55625 #type complete  
ENTRY protein-tyrosine kinase (EC 2.7.1.112),  
TITLE megakaryocyte-associated - human  
ORGANISM #formal name Homo sapiens #common name man  
DATE 24-Feb-1995 #sequence\_revision 24-Feb-1995 #text\_change 07-Jul-1995

ACCESSIONS A55625  
REFERENCE A55625  
#authors Avraham, S.; Jiang, S.; Ota, S.; Fu, Y.; Deng, B.; Dowler, L.L.; White, R.A.; Avraham, H.  
#journal J. Biol. Chem. (1995) 270:1833-1842  
#title Structural and functional studies of the intracellular tyrosine kinase MATK gene and its translated product.

#accession A55625 preliminary; not compared with conceptual translation  
#status #molecule\_type DNA  
#residues 1-507  
CLASSIFICATION #superfamily SH2 homology; protein kinase homology; SH3 homology  
KEYWORDS phosphotransferase  
FEATURE 55-105 #domain SH3 homology #label SH31\

May 31 09:15

US-08-426-509-2.rpr

4

122-211 #domain SH2 homology #label SH2\  
233-485 #domain protein kinase homology #label KIN  
SUMMARY #length 507 #molecular-weight 56469 #checksum 6051

Query Match 100.0%; Score 3727; DB 11; Length 507;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 magrgslvsrafhgdcdaeeiprvsprflrawhppvpsarmptrrwapgtqctkceht 60  
|||||  
Qy 1 MAGRGSLVSRAFHGCDSAEELPRVSPRFLRAWHPPVPSARMPTRRWAPGTQCTKCEHT 60  
|||||

Db 61 rpkpgelafrkgdvvtilaeacnksyrvkhhtsqegllaaagalrerealsadpklslm 120  
|||||  
Qy 61 RPKPGELAFRKGDDVVTILEACENKSWRVKHHTSQEGLLAAGALREREALSADPKLSLM 120  
|||||

Db 121 pwfhgklsqgeavqqlppedglflvresarhpgdyvlcvsfgrdvihyrvlhrdghlti 180  
|||||  
Qy 121 PWFHGKLSQGEAVQQLPPEDGLFLVRESARHPGDYVLCVSFGRDVIHYRVLHRDGHITI 180  
|||||

Db 181 deavffcnlmdmvehyekdkgaictklvprkkghtksaeelaraagwlinlqhltlgaq 240  
|||||  
Qy 181 DEAVFFCNLMDMVEHYSKDKGAICTKLVRPKKHGHTKSAAEELARAGWLINLQHILTLGAQ 240  
|||||

Db 241 igegefagvlqgeylgqkvavknikcdvtaqafldetavmtkmqhenlvrlilghlgaq 300  
|||||  
Qy 241 ICEGEFAGVLOGEYLGQKVAVKNIKCDVTAQAFLDETAVMTKMQHENVRLILGHILGAQ 300  
|||||

Db 301 yivmehvsknlnvflrtgralvntaqlqlfslhvaeqmeyleskklvhrdlaarnilv 360  
|||||  
Qy 301 YIVMEHVSNGNLNVLRTGRALVNTAQLQLFSLHVAEQMEYLESKKLVHRDLAARNILV 360  
|||||

Db 361 sedlvakvsdfglakaerkglslrpkwtapealkhgkftsksdwsgvllwevfsv 420  
|||||  
Qy 361 SEDLVAKVSDFGLAKAERGLDLSRLPKWTAPALKHGKFTSKSDWSGVLLWEVFSY 420  
|||||

Db 421 grapykmslkeveavekyrmeppegcpvhlmscweaeaprrppfrklaeklar 480  
|||||  
Qy 421 GRAPYKMSLKEVSEAVEKYRMEPEGPCPVHVMSSCWEAEAPRRPPFRKLAELKLAR 480  
|||||

Db 481 elrsagapasvsgqdgadgtsprsqep 507  
|||||  
Qy 481 ELRSAGAPASVSGQDAGDSTSPRSQEP 507  
|||||

RESULT 3 A49865 #type complete  
ENTRY protein-tyrosine kinase (EC 2.7.1.112) matk - human  
ALTERNATE\_NAMES megakaryocyte-associated tyrosine kinase  
ORGANISM #formal name Homo sapiens #common name man  
DATE 30-Jun-1995 #sequence\_revision 30-Jun-1995 #text\_change 21-Jul-1995

ACCESSIONS A49865  
REFERENCE A49865  
#authors Bennett, B.-D.; Cowley, S.; Jiang, S.; London, R.; Deng, B.; Grubarek, J.; Grooman, J.E.; Goeddel, D.V.; Avraham, H.  
#journal J. Biol. Chem. (1994) 269:1068-1074  
#title Identification and characterization of a novel tyrosine kinase from megakaryocytes.

#accession A49865 preliminary  
#status #molecule\_type mRNA  
#residues 1-527  
#label BEN  
#cross-references GB:U18974

CLASSIFICATION #superfamily SH2 homology; SH3 homology  
KEYWORDS phosphotransferase

FEATURE  
55-105 #domain SH3 homology #label SH31\  
122-211 #domain SH2 homology #label SH2  
SUMMARY #length 527 #molecular-weight 58473 #checksum 1630

Query Match 91.8%; Score 3423; DB 11; Length 527;  
Best Local Similarity 99.1%; Pred. No. 0.00e+00;  
Matches 462; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Db 1 magrslsvwrfahcdsaeeprvsrflrawhppvpsarmtrwpgtqitkccht 60  
QY 1 MAGRSLSVWRFHCDSAEELPRVSRFLRAWHPPVPSARMTRWPGTQITKCEHT 60  
Db 61 rpkqelafrkdvvtileacenksywrkhhtsqegllaaqalrdgealsadpklslm 120  
QY 61 RPKQELAFKGDVVTILEACENKSWVRKHHTSQEGLLAAQALREREALSADPKLSLM 120  
Db 121 pwfhgkisqgeavqlqppegdglflvresarhpgdyvlcvsfgrdvihrvldrdghlti 180  
QY 121 PWFHGKISQGEAVQQLQPPEDGLFLVRESARHPGDYVLCVSFGRDVIHRVLDRDGHITI 180  
Db 181 deavffcnlmdmvehykdkgaictklivprkthgtkaeeelaragwlnlqhltlgaq 240  
QY 181 DEAVFFCNLMDMVEHYSKDKGAICTKLIVPRKRKHGTSKEELARAGWLNLUHLTLGAQ 240  
Db 241 igegefgavlqgeylgqkvavknikcdvtqaafldetavmtkmqhenlvrlglvllhggll 300  
QY 241 IGEGEFGAVLQGEYLGQKVAVKNIKCDVTAQAFLDETAVMTKMQHENLVRLGLVLLHGGLL 300  
Db 301 yivmehvsknlnvflrtgralvntaqlqfslhvaeqmeyleskklvhrdlaarnilv 360  
QY 301 YIVMEHVSKNLNVFLRTGRALVNTAQLQFSLHVAEQMEYLESKKLIVHRDLAARNILV 360  
Db 361 sedlvakvsdglakaerkldstlrvktapealkhg-fteksdwsfgvllwefvsy 419  
QY 361 SEDLVAKVSDGLAKAERKLDSTLRVKTAPALKHGFTEKSDWSFGVLLWEFVS 420  
Db 420 grapykmslkevsaevekyrmepgecgpgvuhvlnsscweaapp 465  
QY 421 GRAPYKMSLKEVSAEVEKYRMEPEGCGPVUHVLNSSCWEAEP 466

RESULT 4  
ENTRY B55625 #type complete  
TITLE protein-tyrosine kinase (EC 2.7.1.112),  
megakaryocyte-associated - mouse  
ORGANISM #formal name Mus musculus #common name house mouse  
DATE 24-Feb-1995 #sequence\_revision 24-Feb-1995 #text\_change 07-Jul-1995  
ACCESSIONS B55625  
REFERENCE A55625  
#authors Avraham, S.; Jiang, S.; Ota, S.; Fu, Y.; Deng, B.; Dowler, L.L.; White, R.A.; Avraham, H.  
#journal J. Biol. Chem. (1995) 270:1833-1842  
#title Structural and functional studies of the intracellular tyrosine kinase MATK gene and its translated product.  
#accession B55625  
#status preliminary; not compared with conceptual translation  
#molecule type mRNA  
#residues 1-465 #label AVR  
CLASSIFICATION #superfamily SH3 homology; SH2 homology  
KEYWORDS phosphotransferase

FEATURE  
13-63 #domain SH3 homology #label SH31\  
80-169 #domain SH2 homology #label SH2  
SUMMARY #length 465 #molecular-weight 51585 #checksum 6919

Query Match 85.4%; Score 3182; DB 11; Length 465;  
Best Local Similarity 89.7%; Pred. No. 0.00e+00;  
Matches 418; Conservative 37; Mismatches 10; Indels 1; Gaps 1;

Db 1 mptr-wapgtqcmktcensrpkpgelafrkqdmvtileacedksywrakhhgsgqeglla 59  
QY 42 MPTRWAPGTQCITKCEHTRPKPGELAFRKGDVVTILEACENKSWVRKHHTSQEGLLA 101  
Db 60 aaalrqrealstdpklslmpwfhgkisqgeaiaqlqppegdglflvresarhpgdyvlcvsl 119  
QY 102 AGALREREALSADPKLSLMPWFHGKISQGEAVQQLQPPEDGLFLVRESARHPGDYVLCVS 161  
Db 120 fgdrvihrvldrdghltideavvcfnlmdmvehytkdkgaictklivprkkgaksaee 179  
QY 162 FGRDVIHRVLDRDGHLTIDEAVVCNLMDMVEHYSKDKGAICTKLIVPRKRKHGTSKAE 221  
Db 180 elakagwllldqhltdlgaqigegefgavlqgeylgqkvavknikcdvtqaafldetavmt 239  
QY 222 ELARAGWLNLUHLTLGAQIGEGEFGAVLQGEYLGQKVAVKNIKCDVTAQAFLDETAVMT 281  
Db 240 klqhrnlvrlglvllhghlyivmehvsknlnvflrtgralvntaqlqfslhvaeqm 299  
QY 282 KMQHENLVRLGLVLLHGHLYIVMEHVSKNLNVFLRTGRALVNTAQLQFSLHVAEQME 341  
Db 300 yleskklvhrdlaarnilveedlvakvsdglakaerkldstlrvktapealkngfr 359  
QY 342 YLESKKLIVHRDLAARNILVSEDIVAKVSDFGLAKAERKLDSSRLPVKWTAPALKHKG 401  
Db 360 sksdwsfgvllwefvsygrapykmslkevsaevekyrmepgecgpgsvhtlmgscw 419  
QY 402 TSKSDWSFGVLLWEFVSYGRAPYKMSLKEVSAEVEKYRMEPEGCGPVHVLMSSCW 461  
Db 420 eaeparppfrkiveklgrslsvsqspagggaegeaptresqp 465  
QY 462 EAEPARPPFRKLAERLARELSAGAPASVSGQDAGGSTSPRSQEP 507

RESULT 5  
ENTRY A41973 #type fragment  
TITLE protein-tyrosine kinase (EC 2.7.1.112) CSK - chicken (fragment)  
ORGANISM #formal name Gallus gallus #common name chicken  
DATE 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 12-Apr-1995  
ACCESSIONS A41973  
REFERENCE A41973  
#authors Sabe, H.; Knudsen, B.; Okada, M.; Nada, S.; Nakagawa, H.; Hanafusa, H.  
#journal Proc. Natl. Acad. Sci. U.S.A. (1992) 89:2190-2194  
#title Molecular cloning and expression of chicken C-terminal Src kinase: lack of stable association with c-Src protein.  
#cross-references MVID:92196083  
#accession A41973  
#status preliminary  
#molecule type DNA  
#residues 1-450 #label SAB  
#cross-references NCBI:88058; NCBI:88059  
#note sequence extracted from NCBI backbone  
CLASSIFICATION #superfamily protein-tyrosine kinase src; protein kinase

**KEYWORDS**  
 homology; SH2 homology; SH3 homology  
**FEATURE**  
 ATP; phosphotransferase; tyrosine-specific protein kinase  
 16-65  
 82-171  
 193-447  
 201-209  
**SUMMARY**  
 homology; SH2 homology; SH3 homology  
 ATP; phosphotransferase; tyrosine-specific protein kinase  
 #domain SH3 homology #label SH3\  
 #domain SH2 homology #label SH2\  
 #domain protein kinase homology #label KIN\  
 #region protein kinase ATP-binding motif  
 #length 450 #checksum 7

Query Match	50.2%;	Score 1871;	DB 4;	Length 450;
Best Local Similarity	53.8%;	Pred. No. 4.28e-294;		
Matches	235;	Conservative 100;	Mismatches 97;	Indels 5; Gaps 4;
Db	8	wpsgtciacinyfnhtgaeg-dlpfskqdvltavtdkdpwnaknkv-gregipanyv	65	
Qy	47	WAPGTQCTKCE-HTPRPGCEIAPRKGDWTTILECENKSWVRVXHHTSQGEGLLAAGAL	105	
Db	66	qkregvqagikslmpwfhgkitregaellyppetgflflvrestnypgdytlcvscqgk	125	
Qy	106	REREALSDAPKSLMPWFHGKLSQGEAVQOQLPPEDGLFLVRESARHPGDVLCVSGFRD	165	
Db	126	vehyriiysssklsideevyfenlmqlvehytttdadglcslrikpkvmegtvaaqdefsr	185	
Qy	166	VIHYRVLRHDGHLTIDEAVFFCNLMNDMVEHYSKOKGAICTKLVRPKRKHGSAEELAR	225	
Db	186	sgwalnmkdlklqlikgqefgdvmlgdvrygnkvavkcikndataqaflaeasvmtqlrh	245	
Qy	226	AGHLLMLQLTLQAIGEGEFGVLOQYLGQKVAVKNIKCDVTAQAFLEDTAVMTKQGH	285	
Db	246	snlvqlglviveeksglylvtymaksglsvdylrsgvsgvgdcllkfsldvceameyl	305	
Qy	286	ENLVRLGLVLHQ--GLYIVMEHYSKGNLVNFLTRGRALVNTAQIQLQFSLHWAEQGYL	343	
Db	306	eannfvhrdlaarnvlvsedniakvsvdfgltkaeastqdtgkplvkwtapealrekfst	365	
Qy	344	ESKKLVHRDRLAARNVLVSIEDLVAKVSDFGGLAAEKRGKLDSSRLPVKWTAPALKHGKFTS	403	
Db	366	kdsdwsfgillweylsyfgrvpytripkladvprvekgymkdpdgcpaiyvevmkkcwtll	425	
Qy	404	KSDWSFGVLLWEYFSGRAPPYKMSLSKSEAVEKGYRMEPPGCGPVRVILWLSQWEA	463	
Db	426	dpghrpsfhqlreqlvh	442	
Qy	464	EPARPPPFERKIAEKLAR	480	

```

6
RESULT
ENTRY
TITLE
ORGANISM
DATE
22-Nov-1993
#sequence_revision 22-Nov-1993
#text_change 22-Nov-1993
ACCESSIONS
REFERENCE
#authors
#journal
#title
the conserved tyrosine autophosphorylation site (Y416)
of protein-tyrosine kinase (EC 2.7.1.112)
cyl - human
#formal_name Homo sapiens
#common_name man
#cross-references
MUID:92050797
accession
S19024
#status
preliminary
#residues
1-450
#label
PAR

```

[illegible]

```

RESULT 7
ENTRY
TITLE      protein-tyrosine kinase (EC 2.7.1.112) CSK - human
ORGANISM   Homo sapiens
DATE       30-Jun-1992
            #sequence_revision 20-Aug-1994
            #text_change
            05-May-1995
ACCESSIONS JH0559; S38818
REFERENCE   JH0559
            #authors      Braeuninger, A.; Holtrich, U.; Strebhardt, K.;
                        Ruhsamen-Waigmann, H.
            #journal      Gene (1992) 110:205-211
            #title         Isolation and characterization of a human gene that encodes a
                        new subclass of protein tyrosine kinases.
            #cross-references MIMD:92165060
            #accession      JH0559
                        #molecule_type mRNA
                        #residues      1-450
                        #label BRA
            #cross-references EMBL:X59932
            #experimental source lung
            #accession      S38818
REFERENCE   S38818
            #authors      Braeuninger, A.; Karn, T.; Strebhardt, K.;
                        Ruhsamen-Waigmann, H.
            #journal      Oncogene (1993) 8:1365-1369
            #title         Characterization of the human CSK locus.

```



REFERENCE 1 (bases 1 to 1987)  
AUTHORS Bennett,B.D., Cowley,S., Jiang,S., London,R., Deng,B., Grabarek,J.,  
Groopman,J.E., Goeddel,D.V. and Avraham,H.

TITLE Identification and characterization of a novel tyrosine kinase from  
megakaryocytes

JOURNAL J. Biol. Chem. 269 (2), 1068-1074 (1994)

MEDLINE 94117408

COMMENT NCBI gi: 455449

FEATURES Location/Qualifiers

source 1..1987

/organism="Homo sapiens"

/cell\_line="GMK11-5"

/cell\_type="megakaryocyte"

/dev\_stage="adult"

/sequenced\_mol="cDNA to mRNA"

/tissue\_lib="lambda gt11"

<1..63

sig\_peptide

/gene="MATK"

/codon\_start=1

263..1846

/gene="MATK"

/note="NCBI gi: 459550"

/codon\_start=1

/product="tyrosine kinase"

/translation="MAGRSIVSWRAFGHCDASAEILRVSPRFLRAWHPPVPSARMP

RRMAPGTQCIKCEHTRPRPGELAFRRKGDWVTILEACENKSWYRKVHTSGOGLIAA

GALRDGEALSADPKLSIMPFWFGKISQEAHQVQLOPPEDGLFLVRESARHPGDYLCV

SFGRDVHYRLVHDGHLTIDAVFFCNLMDVHERYSKQGAICTKLVRPKRHGTS

AEELARAGMLNLQHLITLGAQIGEGFAGVLQGEYLQGVAVKNIKCDVTRQAFIDE

TAVMTKQWENLVLLGLVTLHQGLYIVMEHVSNGNLVNLTRGRALVNTAQLQFSL

HVAEGMEYLIESKKULVHRDLIARNLIVSEDLIVAKVSDFGLAERKGLDSSRLPVRWTA

PEALKHGFYSKSDWVSGFVLLWEVFSYGRAPYPRKMSKEVSEAEKGYRMEPEGCPG

PVHVLMSQWEAEPPAGHRPSANWPRSWPGSYAVQVPPSQGRTPTVHLAPKPGALTP

PGGFWPQTERVESAMWGH"

568

polya\_site

/gene="MATK"

BASE COUNT 369 a 628 c 672 g 318 t

ORIGIN

Query Match 95.98; Score 1918; DB 50; Length 1987;

Best Local Similarity 99.3%; Pred. No. 0.00e+00;

Matches 1976; Conservative 0; Mismatches 4; Indels 9; Gaps 6;

Db 8 ctgcgtccaaagtgtgcagcgggaccgctcggggtgtgcagccggctcgcgaggccc 67

|||||

Qy 1 CTCGCTCCAAGTTGTGCAGCGCGGACCGCCTCGGGGTGTGCAGCCGGCTCGCGGAGGCC 60

Db 68 tccctggggcgggcgggcggtcggggqcgccccctgagcagaaaacagagaac 127

|||||

Qy 61 TCCTGGGGGGGGGGCGGGGGGGCTCGGGGGGGCCCCCTGAGCAGAAAACAGAGAACC 120

Db 128 aggcctcggtccagtggcaccagctccctacctctgtgccagccgctggcctggca 187

|||||

Qy 121 AGGCTGGGTCCAGTGGCAGCCAGCTCCCTACTCTGTGCCAGCCCGCTGGCGTGTGGCA 180

Db 188 ggcattccacagctcccagctgtgacacattgtcagtgctcgtctcctcagctgcacg 247

|||||

Qy 181 GSCCAATTCCACAGCTCCCGACTGTGACCACTTGCCTCAGTGTGCCTCTCACCTGCCTCAG 240

Db 248 ttctcc-tctggggg-cgatggcgggcgaggtctctcgtttctcgtggggcatttcacg 305

|||||

Qy 241 TTTCCTCTCGGGGGCGATGGCGGGGCGAGGCTCTCTGCTTTCTCTCGGGGCAATTACG 300

Db 306 gctgtgattctgtgaggaactccccgggtgagcccccgcttctcagcagctggcacc 365

|||||

Qy 301 GCTGTGATTCTCTGAGGAACCTTCCC GGGTGAGCCCCGGCTTCCTTCGAGCCTTGGCACC 360

Db 366 cccctcccgctctcagccagatgccaaacagagcgctggcccgagcaccagtgatca 425

|||||

Qy 361 CCCCTCCCGTCTCAGCCAGGATGCCAAAGCAGGCGCTGGCGTTCGCGAAGGGCAGCTGG 420

Db 426 ccaaatcgagcacaccgcccccgaagcaggggagctggccttcgcgaagggcgacgtgg 485

|||||

Qy 421 CCAAATGCGAGCACACCGCGCCCAAGCAGGGAGCTGGCTTCCGCAAGGGCAGCTGG 480

Db 486 tcacattcctggggcctgcgagaaagagctggtaccggtcaagcacacacacagt 545

|||||

Qy 481 TCACCATCTCGAGCGCTCGGAGAACAGAGCTGTACCGCGTCAAGCACACACAGCTG 540

Db 546 gacaggaggggctgctggcagctggggcgctggggcgagggagggcctctccgagacc 605

|||||

Qy 541 GACAGGAGGGGCTGCTGGCAGCTGGGGCGCTGCGGGAGCGGGAGGCCCTCTCCGAGACC 600

Db 606 ccaagctcagcctcatcgctggttccacgggaagatctcgggccagagaggtgtccagc 665

|||||

Qy 601 CCAAGCTCAGCCTCATGCGCTGCTGCCAGCGGAAGATCTCGGGCCAGGAGGCTGTCCAGC 660

Db 666 agctgcagcctcccgagatgggctgtcctggtcggggagtcgcgcgcacccccggcg 725

|||||

Qy 661 AGCTGCAGGCTCCCGAGGATGGGCTGTTCTGTCGGGAGTCGCGCGCCAGCCCGCGCG 720

Db 726 actagctcctgtgggtgagcttggcgcgagctcatccactaccggctgtgtcacccg 785

|||||

Qy 721 ACTAGCTCTGTGCGTGAAGCTTTGGCGGCCAGCGCTCATCCACTACCGCGTGTCTGCACCGCG 780

Db 786 acggccacctcacatcgatgagcgctgtcttctgcaacctcatggacatggtggagc 845

|||||

Qy 781 ACGGCGACCTCACAAATCATGAGGCGCTGTTCTTGCACACCTCATGGACATGGTGGAGC 840

Db 846 attacagaaagacaaggcgctatctgcaccaagctggtgagaccaaagcgaaacacg 905

|||||

Qy 841 ATTACAGAAAGACAAGGGCGCTATCTGCACCAAGCTGCTGACACCAAGCGGAACACG 900

Db 906 ggaaccaagtgcgcgagagagctggccagggcggtggtggttactgaacctgcagcatt 965

|||||

Qy 901 GGACCAAGTCGGCGCAGGAGCTGGCCAGGGCGGTGTTACTGAACTCGACGACTT 960

Db 966 tgacattgggagcacagatcgagagggagagtttgagctgtcctgcaggtgagtagc 1025

|||||

Qy 961 TCACATTGGGACGACACATCGCAGAGGGGAGCTTTGGAGGCTGTCTCGCAGGGTGAGTACC 1020

Db 1026 tggggcaaaagtggccgtgaagaatatcaagtgtgagtgacagcccgagcctcctgg 1085

|||||

Qy 1021 TGGGGCAAAAAGGTGGCGTGAAGAAATATCAAGTGTGTGTCACAGCCCGACGGCTTCTCTGG 1080

Db 1086 acgagacggcgctcatgacagaagatgcaacacgagaacctgggtcgctcctctggggcgtag 1145

|||||

Qy 1081 ACGACGCGCGCTCATGCAAGATGCAACACGAAACCTGCTGCTCTCTGGCGGTGA 1140

Db 1146 tctctgcacagggggctgacattgtcatggagcagctgagcaaggggaacctggtgaaact 1205

|||||

Qy 1141 TCCTGCACCAAGGGGCTACATTGTGTCAGCAGCTGACGAAGGGCAACCTGCTGTAAC 1200

Db 1206 ttctgcggaccggggtcgagcctcgtgacacacgctcagctcctcgagttttctctgc 1265

|||||

Qy 1201 TTTCCTCGCGGAGCCCGGGGTGAGGCCCTCGTGACACCGCTCAGCTCTGACGTTTTCTCTGC 1260

Db 1266 acgtggcgaggggcatggagtagctggagagcaagaagcttgtgcaacgcgagacctggccg 1325

```
Qy 1261 ACCTGGCCGAGGGGATGAGTACCTGAGACGACAGAGCTTTGTGACCGGAGCTGGCGG 1320
Db 1326 cccgaacaatcctggtctcagagacctggtgccaaagtcaggaactttggcctggcca 1385
Qy 1321 CCCGCACATTCCTGCTCTCAGAGGACCTGGTGGCCAAAGGTACCGGACTTTGGCCTGGCCA 1380
Db 1386 aagccagcaggaaggggctagactcaagccgctgcccgctcaagtggacggcgcccgagg 1445
Qy 1381 AAGCCGAGCGGAAGGGGTAGACTCAAGCGCGCTGCCGCTCAAGTGGACGGCGCCGAGG 1440
Db 1446 ctctcaaacacagg---ttcaccaagtcagtgatgctgtgaggtttggggtgctgtct 1502
Qy 1441 CTCTCAACACGGAGGTTCCACACGACGTCGATGCTGGAGTTTGGGGTCTGCTCT 1500
Db 1503 gggaggtcttctcatatggacgggctccgtacctaaaatgtcactgaagaggtgtcgg 1562
Qy 1501 GGGAGGTCTTCTATATGAGCGGGCTCGGTACCCCTAAATGTCTGAAAGAGGTGTCGG 1560
Db 1563 aggcctggagaaaggtaccgcatgaaccccgagggtgtccagcccgctgcacg 1622
Qy 1561 AGGCCGTGGAGAAAGGGTACCGATGCAACCCCGAGGGCTGTCCAGGCCCGCTGCACG 1620
Db 1623 tctcatgagcagctgctgggagcagagcc-gcccgccggccacccttcgcaaatgg 1681
Qy 1621 TCCTCATGACGAGCTGCTGGAGGCGAGCGCCGCCGCCGACCCCTTCGCAAACTGG 1680
Db 1682 ccgagaagctggccgggagctacgagtgaggtgccccagcctccgtctcagggcagg 1741
Qy 1681 CGGAGAGCTGGCCGGGAGCTACGAGTGCAGGTGCCCGAGCTCCCTCTCAGGCGAGG 1740
Db 1742 acgcccagcg-tccacctgccccgaagcaggagcctgacccaccggttg--ccct 1798
Qy 1741 AGCCGAGCGGCTTCACCTGCCGCCGAGACGAGGACCTTCGCCGCCCGGCTGGGGCCCT 1800
Db 1799 tggcccccagagccagagatggaagtgccggtgcccgtgggggacctgacagcccaagg 1858
Qy 1801 TGGCCCCAGAGCCGAGAGTGGAGTGGCGGTGGGGGCTGACGAGGCCCAAG 1860
Db 1859 aggtccagcgcggaagtcactctctggtgccacagcagggtggcccaagtg 1918
Qy 1861 AGGGTCCAGCGGGCGAAGTCATCTCTCTGTCGCCACAGCGGGCTGGCCACGCTAGGG 1920
Db 1919 ggtctggcgcccggtggacacccagacctggcaaggatgatgcccgataaagcgg 1978
Qy 1921 GGCTCTGGCGGGCGGTGGACACCCCGAGCTGCCAGAGGTATGATCCGCCATTAAGCGG 1980
Db 1979 attctaagg 1987
Qy 1981 ATTCTAAGG 1989
```

```
RESULT 2
LOCUS HSHYLTk 1968 bp RNA PRI 10-OCT-1994
DEFINITION H.sapiens HYL tyrosine kinase mRNA.
ACCESSION X77278
KEYWORDS HYLTK gene; nonreceptor protein tyrosine kinase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Carnivorini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 1968)
AUTHORS Sakano,S., Iwama,A., Inazawa,J., Ariyama,T., Ohno,M. and Suda,T.
TITLE Molecular cloning of a novel non-receptor tyrosine kinase, HYL
(hematopoietic consensus tyrosine-lacking kinase)
JOURNAL Oncogene 9 (4), 1155-1161 (1994)
MEDLINE 94181267
REFERENCE 2 (bases 1 to 1968)
AUTHORS Iwama,A.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-1994) to the EMBL/GenBank/DBJ databases. A.
Genetics, Kumamoto University School of Medicine, 2-2-1 Honjo,
Kumamoto 860, JAPAN
NCBI gi: 471312
FEATURES
source Location/Qualifiers
1..1968
/organism="Homo sapiens"
/tissue_type="hematopoietic cell"
/clone_lib="UT-7 cDNA library"
/chromosome="19p13"
1..1968
/genes="HSHYLTk"
208..1731
/genes="HSHYLTk"
/notes="NCBI gi: 557272"
/codon_start=1
/product="HYL tyrosine kinase"
/translation="MAGRSLVSWRAPHGCDSEAEIPRSPFLRAHPPVPSARMP
RRWAPGTCTKCHTRPKPGEIAPRKGQVWVILEACENKSHYRVKHTTSGEGILAA
CALREALSADPKLSIMFWHFKISQEAQQVQLPPEDGLFLVRESARHPDYVLGV
SFGDVIHYRVLDHGLTIDEAVFFCNLMDVHYSKDKGAICTKLVRPRKHGTS
AEELARAGMLINLQHLITLQAIQEGEGAVLQGEYLGQKAVKNIKDVTQAFLDE
TAVWTKQHENLVPLLVILVILHQLYIVMEHVSKNLVNLFTRGRALVNTQILQFSL
HVAEGMEYLESKLVLRDIAARNILVSDILVAKVSDFLAKAERKGLDSSRLPVKMTA
PEALKHGKTSKSDVMSFGLIWEVFSYGRAPYPKMSLKEVSEAVEKGRMPEPGCP
GPVHVLSSQWEAEAPARPPFRKLAELRLSAGAPASVSGQADUGSTSPRSQEP"
BASE COUNT 391 a 614 c 652 g 311 t
ORIGIN
Query Match 95.7%; Score 1914; DB 45; Length 1968;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 1945; Conservative 0; Mismatches 1; Indels 5; Gaps 3;
Db 1 cggagggcctcctggggcgggcgggcgggcggtcgggggcgccctcgagcagaaa 60
Qy 52 CGGAGGCCCTCTTGGGGCGGGCGGGGGG--GCTGGGGGGCGCCCTGACGAGAAA 109
Db 61 caggaagaacacaggctcgtccagtggcaccagctccctacctctgtgcagcgact 120
Qy 110 CAGGAAGAACACAGGCTCGGTCCAGTGGCACCAGCTCCCTACCTCTGTGCCAGCGCT 169
Db 121 ggcctgtggcagggcattcccagaggtcccagactgtgaccacttgtcagtgctctc 180
Qy 170 GGCCTGTGGCAGGCCATTCCACGCTCCCGACTGTGACCACTTGTCTCAGTGTGCTCTC 229
Db 181 acctgcctcagtttccctctggggg--cgatggcgggcgagggctctctggttctcg 239
Qy 230 ACCTGCCTCAGTTTCCCTCTGGGGGGCGATGGGGGGGAGGGCTCTCTGTTTCTCTGG 289
Db 240 ggcatttcagcgtgtgattctgtgaggaacttcccgggtgagcccccgcttctccg 299
Qy 290 GGCATTTACGGCTGTGATTCGTGAGGAACTTCCCGGGGTGAGCCCCCGCTTCTCTCG 349
Db 300 agcctggcaccctccctccgtctcagccagagatgcaacgagggcgctggccccggcac 359
```

```
REFERENCE 1 (bases 1 to 1968)
AUTHORS Sakano,S., Iwama,A., Inazawa,J., Ariyama,T., Ohno,M. and Suda,T.
TITLE Molecular cloning of a novel non-receptor tyrosine kinase, HYL
(hematopoietic consensus tyrosine-lacking kinase)
JOURNAL Oncogene 9 (4), 1155-1161 (1994)
MEDLINE 94181267
REFERENCE 2 (bases 1 to 1968)
AUTHORS Iwama,A.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-1994) to the EMBL/GenBank/DBJ databases. A.
Genetics, Kumamoto University School of Medicine, 2-2-1 Honjo,
Kumamoto 860, JAPAN
NCBI gi: 471312
FEATURES
source Location/Qualifiers
1..1968
/organism="Homo sapiens"
/tissue_type="hematopoietic cell"
/clone_lib="UT-7 cDNA library"
/chromosome="19p13"
1..1968
/genes="HSHYLTk"
208..1731
/genes="HSHYLTk"
/notes="NCBI gi: 557272"
/codon_start=1
/product="HYL tyrosine kinase"
/translation="MAGRSLVSWRAPHGCDSEAEIPRSPFLRAHPPVPSARMP
RRWAPGTCTKCHTRPKPGEIAPRKGQVWVILEACENKSHYRVKHTTSGEGILAA
CALREALSADPKLSIMFWHFKISQEAQQVQLPPEDGLFLVRESARHPDYVLGV
SFGDVIHYRVLDHGLTIDEAVFFCNLMDVHYSKDKGAICTKLVRPRKHGTS
AEELARAGMLINLQHLITLQAIQEGEGAVLQGEYLGQKAVKNIKDVTQAFLDE
TAVWTKQHENLVPLLVILVILHQLYIVMEHVSKNLVNLFTRGRALVNTQILQFSL
HVAEGMEYLESKLVLRDIAARNILVSDILVAKVSDFLAKAERKGLDSSRLPVKMTA
PEALKHGKTSKSDVMSFGLIWEVFSYGRAPYPKMSLKEVSEAVEKGRMPEPGCP
GPVHVLSSQWEAEAPARPPFRKLAELRLSAGAPASVSGQADUGSTSPRSQEP"
BASE COUNT 391 a 614 c 652 g 311 t
ORIGIN
```

|||||  
QY 350 AGCTGGACCCCTCCGCTCTAGCCAGGATGCAACAGAGCGCTGGGCCCGGGAC 409  
Db 360 ccagtgtatcaccaaatcgagcacaccccccacagccaggagagctggcttcgcgaa 419  
QY 410 CCAGTGTATCACCAAATGCGAGCACACCGCCCAAGCCAGCGGAGCTGGCCTTCGGAA 469  
Db 420 gggcgagctggtcaccatcctggagcctcgagagcaaaagagctggtaccgctcaagca 479  
QY 470 GGGCGAGCTGGTACCATCTGTAGGCGCTCGGAGAACAGAGCTGTACTCCGCTCAAGCA 529  
Db 480 ccacaccagtggacaggagggtctgycagctggggcgctgcggagcgaggccct 539  
QY 530 CCACACCAGTGGACAGGAGGGGCTCTGGCAGCTGGGGCGCTCGCGGAGCGGAGGCCCT 589  
Db 540 ctccqcaqaccccaagctcagcctcatgccgtggttccacgggaagatctcgggccagga 599  
QY 590 CTCGCGAGACCCCAAGCTCAGCCTCATGCCGTGGTTCCACGGGAAGATCGCGGGCAGGA 649  
Db 600 ggctgtccagcagctgcagcctccgaggatgggctgttctcgtggtgcggagctccgacg 659  
QY 650 GGCTGTCCAGCAGCTGCAGCTCCGAGGATGGGCTGTCTGTGGGAGTCCGCGCG 709  
Db 660 ccaaccgcgcgactacgtcctgctgagcttgagcttggccgcagcgtcatcactaccgct 719  
QY 710 CCACCGCGCGACTACGTCTGTGGCTCAGCTTTGGCGGGCAGCTCATCCACTACCGGT 769  
Db 720 gctgcacgcgcagggccacctcacatcgatgagccgctgttcttctgaacctcatgga 779  
QY 770 GTTGCAACCGGAGCGGCACCTCACAATCGATGAGCGCGCTGTTCTTCTGCAACCTCATGA 829  
Db 780 catggtgagctaccagcaaggaagggcgctatctgcacaaagctggtgagacaaa 839  
QY 830 CATGCTGGAGCATTACAGCAAGCAAGGGCGCTATCTGCACCAAGCTGGTGAGACCAA 889  
Db 840 gcggaaacagggaccaagtcggccgagagagctggccagggcgggctggtactgaa 899  
QY 890 CGGAAACACGGGACAAAGTCGGCGCAGGAGGCTGGCCAGGGCGGCTGCTTACTGAA 949  
Db 900 cctgcagcatttgacattgggagcagatcggagagggagagtttgagactgctctgca 959  
QY 950 CCTGCAGCATTTGACATTGGGACACAGATCGGAGAGGAGATTTGGAGCTCTCTGCA 1009  
Db 960 gggtagtacctggggcaaaagtgccgtgaagaatatcaagtgtgatgacagccca 1019  
QY 1010 GGGTCACTACCTGGGGCAAAAGTGGCGCTGAGAAATATCAAGTGTGATGACAGCCCA 1069  
Db 1020 ggcttctggagcagcgccgtcatgcgaagatgcaacacgagaaacctggtgcgtct 1079  
QY 1070 GGCCTCTCTGGACGAGAGCGCCGTCTAGACGAAGATGCACAGAGAACCTGGTGGCT 1129  
Db 1080 cctggcgctgatcctgaccagggctgtacattgtcatggagcagctgagcaaggcaa 1139  
QY 1130 CCTGGCGGTGATCTGCAACAGGGGCTGTACATTGTATGGAGCACTGTGAGCAAGGGAA 1189  
Db 1140 cctggtgaactttctgcgacccggggtcgagccctcgtgaacacgcctcagctcctgca 1199  
QY 1190 CTTGGTGAACCTTTCTGGGACCGGGGTGAGCCCTCTGTGAACACCGCTCAGCTCTGCA 1249  
Db 1200 gttttctctgcagctggccgagggcagtagtacctggagagcaagaagcttgtgcacc 1259  
QY 1250 GTTTTCTCTGCAGCTGGCCGAGGGATGGAGTACCTGGAGCAGAGAGCTTGTGACCG 1309  
Db 1260 cgaacctggccgcccgaacatcctggtctcagaggacctggtggccaaggtcagcgactt 1319

|||||  
QY 1310 CGACCTGGCCGCCCAACATCTCTGCTCAGAGGACCTGTGTGGCAAGGTCAAGCACTT 1369  
Db 1320 tggcctggccaaagccagcgaggaaggggtctagactcaagcgcgctgcccgtcaagtggac 1379  
QY 1370 TGGCCTGGCCAAAGCCGAGGAGGGGCTAGACTCAAGCGGCTGCCCGTCAAGTGCAC 1429  
Db 1380 ggcgcgcgaggtctcacaacacgggaagttcaccagcaagtcggtctctggagtttgg 1439  
QY 1430 GGCGCCCGAGGCTCTCAACACAGGGAAGTTCCACAGCAAGTCTGCGATGCTCGAGTTTG 1489  
Db 1440 gdtgctctgggaggtcttctcatatgagcgggtccgtaccctaaaatgtcaactgaa 1499  
QY 1490 GGTGCTGCTCTGGAGGTCTTCTCATATGAGAGGGCTCCGTACCTAAATATGCTAC 1549  
Db 1500 agagtgctcggaggcgtggagaaggggtaccgcatgaaccccccgagggtgtccagg 1559  
QY 1550 AGAGTGTCCGAGGGCTTCCAGAGGGGTACCGCATGGAACCCCGAGGGCTGTCCAGG 1609  
Db 1560 cccgtgcaagctcctcatgagcagctgctgggagcagagcccgcccgccgacacctt 1619  
QY 1610 CCCCCTGCAGCTCTCATGAGCAGCTGCTGGAGGACAGCCCGCCGCGGCAACCTT 1669  
Db 1620 ccycaaactggccgagaagctggccgggagctacgcagtgcaggtgccccagctccgt 1679  
QY 1670 CGCGAACTGGCGAGAGCTGGCCGGGAGCTACGCACTGCAAGTGGCCAGCCCTCGCT 1729  
Db 1680 ctcaggcagcagcgagcgtctccacctgcgccccgaagcagagcctgaccccaacc 1739  
QY 1730 CTCAGGCGAGGAGCGCGCGCTCCACCTCGCCCCGAAGCAGAGGAGCCCTGACCCACCC 1789  
Db 1740 ggtgg--cccttggccccagagcagagagagtgagagtgcgcggtggggcactgac 1797  
QY 1790 GGTGGGGCCCTTGGCCCCACAGGACCGAGAGAGTGGAGCTGGCGCGTGGGGGCACTGAC 1849  
Db 1798 caggcccaagagggtccagcggggcaagctcatcctcctggtgcccaagcagggtcgg 1857  
QY 1850 CAGGCCCAAGAGGGTCCAGCGGCGCAAGTCACTCCTGTGTGCCCAACAGCAGGGCTGG 1909  
Db 1858 cccaogtagggggtctggcgggccgctggacaccccgacacctcggaaggtgatcgccc 1917  
QY 1910 CCCAGTAGGGGGCTCTGGCGGCGCGTGGACACCCGAGACCTTGCAGAGGATGATCGGCC 1969  
Db 1918 gataaagcaggtattcgaaggtcctaaaaa 1948  
QY 1970 GATAAAGACGGATTCTAAGGACTTAATAAAA 2000  
RESULT 3  
LOCUS MUSNTK 1734 bp mRNA ROD 30-JUN-1994  
DEFINITION Mus musculus tyrosine protein kinase (Ntk) mRNA, complete cds.  
ACCESSION L27738  
KEYWORDS tyrosine protein kinase.  
SOURCE Mus musculus (strain BALB/c, sub\_species domesticus) day 16 fetus  
thymus cDNA to mRNA.  
ORGANISM Mus musculus  
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;  
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.  
REFERENCE 1 (bases 1 to 1734)  
AUTHORS Chow, L.M.L., Jarvis, C.D., Hu, Q., Nye, S.H., Gervais, F.G.,  
Veillette, A. and Matis, L.A.  
TITLE Ntk: A Csk-related protein-tyrosine kinase expressed in brain and T  
lymphocytes  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91, 4975-4979 (1994)



\*\*\*\*\*

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
	1	1906	95.3	1942	15	Q84888	DNA encoding cytoplasm	0.00e+00
	2	114	5.7	147	8	Q49748	PTK gene SAL-D4 parti	4.14e-50
	3	114	5.7	147	16	T03091	Protein tyrosine-kin	4.14e-50
	4	94	4.7	1047	2	Q10572	Human Natriuretic Pep	3.58e-38
	5	81	4.1	1047	2	Q10572	Human Natriuretic Pep	3.58e-38
	6	78	3.9	1611	7	Q46688	Human Natriuretic Pep	1.50e-30
	7	68	3.4	1254	3	Q13983	Human pp60 c-src gene	8.26e-29
	8	68	3.4	1602	7	Q46687	Lck gene fused with p	4.47e-23
	9	65	3.2	1739	4	Q27539	Chicken pp60 c-src ge	4.47e-23
	10	65	3.2	1739	4	Q27539	Tyrosine Kinase recep	2.23e-21

## ALIGNMENTS

RESULT	1
ID	Q84888 standard; cDNA to mRNA; 1942 BP.
AC	Q84888;
DT	26-OCT-1995 (first entry)
DE	DNA encoding cytoplasmic tyrosine kinase.
KW	cytoplasmic; tyrosine kinase; blood; cell differentiation;
KW	screening; anticancer agent; ds.
OS	Homo sapiens.
HFH	Key
FT	Location/Qualifiers
FT	CDS 208..1731
FT	/*tag= a
FT	/product= cytoplasmic_tyrosine_kinase
FT	/note= "see R71133"
FT	misc feature 349..540
FT	/*tag= b
FT	/note= "encodes SH3 domain (see R71129)"
FT	misc feature 571..795
FT	/*tag= c
FT	/note= "encodes SH2 domain (see R71130)"
FT	misc feature 904..1641
FT	/*tag= d
FT	/note= "encodes tyrosine kinase domain (see R71131)"
FT	misc feature 331..1728



FT /\*taq= e  
FT /note= "encodes N-terminally truncated form  
FT of the enzyme (see R71132)"  
PN W09506113-A.  
PD 02-MAR-1995.  
PF 25-AUG-1994; J01411.  
PR 25-AUG-1993; JP-210403.  
PR 29-MAR-1994; JP-058553.  
PA (ASAH ) ASAH KASEI KOGYO KK.  
PI Sakano S;  
DR WPI: 95-106842/14.  
DR P-PSDB; R71129-33.  
PT Cytoplasmic tyrosine kinase and antibody recognising it - for  
PT screening chemical substances for tyrosine kinase inhibitory or  
PT activating activity for use as cancer therapy  
PS Claim 7; Page 49-50; 58pp; English.  
CC This DNA encodes a cytoplasmic tyrosine kinase which has enhanced  
CC expression in connection with blood cell differentiation. It was  
CC isolated from the human UT-7 blood cell line. The DNA sequences and  
CC antibodies raised against the enzyme, are useful for screening agents  
CC for inhibiting or activating activity on the tyrosine kinase, for  
CC use as anticancer agents.  
SQ Sequence 1942 BP; 365 A; 615 C; 651 G; 311 T;

Query Match 95.3%; Score 1906; DB 15; Length 1942;  
Best Local Similarity 99.6%; Pred. No. 0.00e+00;  
Matches 1938; Conservative 0; Mismatches 2; Indels 5; Gaps 3;  
DB 1 cggaggccctcctggggggcgcgcgcgctcgggggcgccccctgagcagaaaa 60  
|||||  
QY 52 CGGAGGCCCTCCTGGGGGGCGCGCGGGCG--GCTGGGGGGCGGCCCTGAGCGA 109  
|||||  
DB 61 caggaaacaccaggtcgtccagtggcaccagctccctacctctgtgccagcgact 120  
|||||  
QY 110 CAGGAAGAACAGGCTCGGTCCAGTGGGACCCAGCTCCCTACCTCTGCGCAGCGCTC 169  
|||||

DB 121 ggcctgtggcaggccattcccgagctcccgcagctgcccagctgtgacctgtcagtgctctc 180  
|||||  
QY 170 GGCCTGTGGCAGGCCATTCCAGCTCCCGACTGTGACCACTTGTCACTGTGCTCTC 229  
|||||  
DB 181 acctgctcagtttccctctggggg-cgatggggggcgaggtctctgtttcctggcg 239  
|||||  
QY 230 ACTTGCCTCAGTTTCCCTCTGGGGGGCGATGGGGGGGAGGCTCTGTGTTTCTGGCG 289  
|||||  
DB 240 ggcatttcacggctgtgattctgtgaggaaattccccgggtgagcccccgcttctccg 299  
|||||  
QY 290 GGCATTACGGCTGTGATTCTGTGAGGAACCTCCCGGGGTGAGCCCCCGCTTCTCCG 349  
|||||  
DB 300 agcctggcaccctcccgctctcagcagagatgccaaagggcgtggccccgggac 359  
|||||  
QY 350 AGCCTTGGCACCCCTCCCGTCTCAGCCAGGATGCCAACAGGCGCTGGGCGCCGGGCAC 409  
|||||  
DB 360 ccagtgatacaaaaatgcagacacaccccccagcgagggtggccttcgcgaa 419  
|||||  
QY 410 CCAAGTGATACCAAAATGCGAGCACACCCGCCCAAGCCAGGGGAGCTGGCCTTCGGCA 469  
|||||  
DB 420 gggcgacgtggtcaacatcctggagcctcgacacaaagactggtaccgctcaagca 479  
|||||  
QY 470 GGGCGAGCTGCTACCATCTCGAGGGCTCGGAGAACAGAGCTGTGACCGCTCAAGCA 529  
|||||  
DB 480 ccacaccagtggacagggggtcgtggcagctggggcgctggggcgaggggcccc 539  
|||||  
QY 530 CCACACCACTGGACAGGAGGGGCTCTGGCAGCTGGGGGCGCTGCGGAGCGGGGCCCT 589  
|||||

DB 540 ctccgagaccccaagctcagcctcatgccgtgggtccacgggaagatctcgggccagga 599  
|||||  
QY 590 CTCGCGACACCCCAAGCTCAGCCTCATGCCGTGGTTCCACGGGAAGATCTCGGGCGAGGA 649  
|||||  
DB 600 ggctgtccagcagctgcagcctcccagagatgggctgttctcgtgtgcgggagtcgcgcgcg 659  
|||||  
QY 650 GGCTGTCCAGCAGCTGCAGCCTCCGAGGATGGGCTGTTCTGTCGGGAGTCCGCGCG 709  
|||||  
DB 660 ccacccggcgactacgtcctgtgcgtgagctttggcgcgagctcatccactaccggt 719  
|||||  
QY 710 CACCCCGGGGACTTACGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769  
|||||  
DB 720 gctgcacgcgagcgccacctcacatcgtgagcgctgttcttctgcaacctcatgga 779  
|||||  
QY 770 GCTGCACCGCGAGCGGCCACCTCACATCGATGAGGCCGTGTTCTTCGACACCTCATGGA 829  
|||||  
DB 780 catggtgagcattacagcaagacaaggcgctatctgcaccaagctggtgagaccaa 839  
|||||  
QY 830 CATGCTGGACATTACAGCAAGCAAGAGGGCGCTATCTGCACCAAGCTGCTGACACCAA 889  
|||||  
DB 840 ggcgaacacgggaccaaagtcggcgagggagctggccaggggcggttactgaa 899  
|||||  
QY 890 GCGGAACACGGGACCAAGTCCGCCGAGGAGAGCTGGCCAGGGCGGCTGTTACTGAA 949  
|||||  
DB 900 cctgcagcatttgacattgggagcacagatcggagggagagtttgagctgctctgca 959  
|||||  
QY 950 CTGACGACATTGACATTGGGAGCACAGATCGGAGGGGAGTTTGGAGCTGCTCTGCA 1009  
|||||  
DB 960 gggtgagtacctggcgcaaaagtgccgtgaagatatacaagtgtgatgtgacagccca 1019  
|||||  
QY 1010 GGGTCAGTACTGGGGGCAAAAGGTGGCCGTGAGGAATATCAAGTGTGATGTGACACCCA 1069  
|||||  
DB 1020 ggcctcctggacgagcggcgctcatgacgaagatgcaacacgagaacctgtgcgtct 1079  
|||||  
QY 1070 GGCCTTCTGACGACGACGGCCCTCATGACGAAGATGCCAACACGAGAACCCTGGTGGCT 1129  
|||||  
DB 1080 cctgggctgactcctgcaccaggggtgtacattgtcagagcacgtgagcaagggcaa 1139  
|||||  
QY 1130 CTGSGGCTGATCTTCGACCCAGGGCTGTACATTGTCTGATGACGACGTGACGAAGGCA 1189  
|||||  
DB 1140 cctggtgaaatttctgcggaccggggtcgagccctcgtgaacaccgctcagctctgca 1199  
|||||  
QY 1190 CTTGTTGAACTTTCTGGCGACCGGGGTGCGGCCCTCTGACACCGCTCAGCTCTGCA 1249  
|||||  
DB 1200 gttttctctgcagtgcccgagggcagtgagtaacctggagagcaagaagttgtgcaaccg 1259  
|||||  
QY 1250 GTTTCCTCTGCGAGTGGCGCGAGGCGATGGAGTACCTGGAGACGAAGAAGCTTGTGACCG 1309  
|||||  
DB 1260 cgaactggcgcccgcaaacactcctggtctcagagagcagctggtggcaaggtcagcagctt 1319  
|||||  
QY 1310 CGACTTGGCGCCGCCAACATCTCTGCTCAGAGGACCTGGTGGCCAAAGGTCTGACGACTT 1369  
|||||  
DB 1320 tggcctggccaaagccgagcggaagggttagactcaagccgggtgcccgtcaagtggac 1379  
|||||  
QY 1370 TGGCCTGGCCAAAGCGGAGGGGCTAGACTCAAGCGGCTGCCCGCTCAAGTGGAC 1429  
|||||  
DB 1380 ggcgcccgaggtctcacaacgggaaagttcaccagcaagtcggagtgctggagtttgg 1439  
|||||  
QY 1430 GGGCGCCGAGGCTCTCAACACGGGAGTTTACCAGAGCAAGTCCGATGCTCTGGAGTTTGG 1489  
|||||  
DB 1440 ggtgctgctgggaggtcttctcatatggagcggtccgtaccctaaaaatgtcactgaa 1499  
|||||  
QY 1490 GGTGCTCTCTGGAGGCTCTTCTCATATGACGCGGCTCCGTACCCCTAAATGTTACTGAA 1549  
|||||

Db 1500 agagtgctcgagggcgtggagaagggtaccqcatgaacccccaggggtgtccagg 1559  
|||||  
Qy 1550 AGAGGTGTGGAGGGCCGTGCAGAACGGGTACCGCATGCAACCCCGGAGGGCTGTCCAGG 1609  
Db 1560 cccgtgcaagtcctcatgagcagctgctgggagcagagcccgccgcgcagccaccctt 1619  
|||||  
Qy 1610 CCCCCTGCAGCTCTCATGACGAGCTGCTGGCAGGCAGACCCCGCCCGCCGCCACCCCTT 1669  
Db 1620 ccgcaaaactggccgagaagctggcccgagggtacgcagtgcaagtgccccagctccgt 1679  
|||||  
Qy 1670 CGGCAAACTGGCCGAGAGCTGGCCGGGAGCTACGCACTGCAGTGCAGCCACCTTCGT 1729  
Db 1680 ctcaaggcagcagcgcagcgtctccacctcgccccgaagccagcagcgtgaccccaacc 1739  
|||||  
Qy 1730 CTCAGGCGAGAGCGCCGACGGCTCCACTCGCCCGGAGCCAGGAGCCCTGACCCACCC 1789  
Db 1740 ggtgg--ccctggcccaagcagcagagagagtgagagtgagcgtggcgctggggcactgac 1797  
|||||  
Qy 1790 GGTGGGCCCTTTGGCCCGCAGAGACCCAGAGAGTGGAGAGTGGCGCTGGGGCAGCTGAC 1849  
Db 1798 caggcccaaggaggtccaggcgggcagtcctcctcctggtgcccaagcaggggtgg 1857  
|||||  
Qy 1850 CAGGCCCAAGGAGGTCAGGCGGGCAAGTCATCTCTGCTGCCACAGCAGGGGGTGG 1909  
Db 1858 cccacgtagggggtctggcgggccgtggacccccagacctgcgaaggtgatcgccc 1917  
|||||  
Qy 1910 CCCAGGTAGGGGGCTCTGGCGGGCCGCTGCACACCCAGACCTGGCAGGATCATCGCCC 1969  
Db 1918 gataaagcaggtattcgaagctct 1942  
|||||  
Qy 1970 GATTAAGCGGATCTTAGGACTCT 1994

RESULT 2

ID Q49748 standard; DNA; 147 BP.  
AC Q49748;  
DT 10-MAR-1994 (first entry)  
DE PTK gene SAL-D4 partial sequence.  
KW PTK; protein tyrosine kinase; catalytic domain; c-kit; FGF;  
KW fibroblast growth factor receptor; megakaryocyte; amplification;  
KW primer; polymerase chain reaction; PCR; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 2..147  
FT /tag= a  
FT misc\_feature 1..21  
FT /tag= b  
FT /note= "pTK1/3 primers"  
FT misc\_feature 122..147  
FT /tag= c  
FT /note= "pTKKW primer"  
PN W09315201-A.  
PD 05-AUG-1993.  
PR 22-JAN-1992; US-826935.  
PA (NENE-) NEW ENGLAND DEACONESS HOSPITAL.  
PI Avraham H, Cowley S, Groopman J, Scadden D;  
DR P-PSDB; R41895.  
PT New protein tyrosine kinase genes and proteins encoded by genes -  
PT are of human mega-karyocytic origin  
PS Claim 2; Fig 2; 60pp; English.  
CC PTK genes were identified using two sets of degenerative  
CC oligonucleotide primers: a first set which amplifies all PTK DNA

CC segments (Q49743-44), and a second set which amplifies highly  
CC conserved sequences present in the catalytic domain of the c-kit  
CC subgroup of PTKs (Q49745-46). The PTK genes identified are described  
CC in Q49747-57 and R41897-02.  
CC SAL-D4 is expressed in several megakaryocytic cell lines, but not  
CC in erythroid cell lines. The SAL-D4 expression prod. exhibited  
CC significant sequence homology with known protein tyrosine kinases  
CC of the FGF receptor family.  
SQ Sequence 147 BP; 37 A; 44 C; 42 G; 24 T;  
Query Match 5.7%; Score 114; DB 8; Length 147;  
Best Local Similarity 92.5%; Pred. No. 4.14e-50;  
Matches 124; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Db 10 cacagagacctagcagcagcagcaacatctgtctcagagacctgtgaaccaaggtcagc 69  
|||||  
Qy 1305 CACGCCGACCTGGCGCCGCCCAACATCTCTGCTCAGAGGACCTGGTGGCCAGGTCAGC 1364  
Db 70 gactttgctgtgccaagccgagcgaggggctagactcaagccggtgccccgtcaaa 129  
|||||  
Qy 1365 GACTTTGGCTGGCCCAAGCCGAGCGAGGGGCTAGACTCAAGCCGGCTGCCCGTCAAG 1424  
Db 130 tggatggtcccgca 143  
|||||  
Qy 1425 TGGACGGCGCCCGCA 1438

RESULT 3

ID T03091 standard; DNA; 147 BP.  
AC T03091;  
DT 14-FEB-1996 (first entry)  
DE Protein tyrosine-kinase SAL-D4 DNA fragment.  
KW Protein tyrosine-kinase; PTK; SAL-D4; agonist; cell growth;  
KW differentiation; ss.  
OS Homo sapiens.  
PN W09527061-A1.  
PD 12-OCT-1995.  
PF 04-APR-1995; U04228.  
PR 04-APR-1994; US-222616.  
PA (GETH) GENENTECH INC.  
PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;  
PI Wood WI;  
DR WPI; 95-366160/47.  
DR P-PSDB; R85923.  
PT Agonist antibodies which activate specific protein tyrosine  
PT kinase(s) - also activate chimeric proteins of kinase extracellular  
PT domain and Ig constant domain, useful for studying, and therapeutic  
PT modulation of, cell growth and differentiation  
PS Disclosure; Page 35-36; 125pp; English.  
CC DNA probes based on protein tyrosine-kinase (PTK) sequences were used  
CC to screen cDNA libraries to identify novel PTK genes. A SAL-D4 gene  
CC fragment (T03091) was isolated from several megakaryocytic cell line  
CC libraries and encoded a protein (R85923) related to the CSK family  
CC of intracellular PTKs. The gene fragment can be used to identify  
CC other new PTK genes, or to design drugs, peptides or antisense  
CC constructs that modulate PTK activity.  
SQ Sequence 147 BP; 37 A; 44 C; 42 G; 24 T;  
Query Match 5.7%; Score 114; DB 16; Length 147;  
Best Local Similarity 92.5%; Pred. No. 4.14e-50;  
Matches 124; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Db 10 cacagagacctagcagcagcagcaacatctgtctcagagacctgtgaaccaaggtcagc 69  
|||||

